IMPERIAL COLLEGE, UNIVERSITY OF LONDON

Study of $|V_{td}/V_{ts}|$ Using a Sum of Exclusive $B \to X\gamma$ Final States Reconstructed with the *BABAR* Detector

by

Mark James Tibbetts

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

in the
School of Natural Sciences
Department of Physics

June 2010

Abstract

Experimental data collected with the *BABAR* detector consisting of 470.9±2.8 million $B\overline{B}$ events are used to measure the sum of seven exclusive $B \to X_{d(s)}\gamma$ transitions, where $X_{d(s)}$ is any non-strange (strange) charmless hadronic state. For each transition flavour, measurements are made in the hadronic mass ranges $0.5 \leq m_X < 1.0\,\mathrm{GeV}/c^2$ and $1.0 \leq m_X \leq 2.0\,\mathrm{GeV}/c^2$. These are extrapolated and combined in a model-dependent way to obtain the ratio of the total branching fractions, $\mathcal{B}(B \to X_d\gamma)/\mathcal{B}(B \to X_s\gamma) = 0.0456 \pm 0.0110 \pm 0.0097$ where the first error is statistical and the second error is systematic. This is interpreted as a measurement of the ratio of CKM matrix elements $|V_{td}/V_{ts}| = 0.211 \pm 0.023 \pm 0.022 \pm 0.001$ where the final error is due to theoretical uncertainty.

Acknowledgements

Firstly I thank my supervisor, Paul Dauncey. Under his watchful eye I have matured in the last four years from a new graduate with a talent for parroting answers in exams to something more akin to a real scientific researcher. His patience with me during this transition cannot be understated. Neither can the importance of his clarity of thinking and direct critiques while discussing work I have presented to him; even when separated by eight time zones. Paul is an excellent supervisor and I am honoured to have worked with him.

The work presented in these pages would not exist without the committed work of Debbie Bard. Her dedication to this analysis before my joining the *BABAR* Collaboration and her valuable assistance after I took on the task of updating her work has been essential to its successful completion. It has been a pleasure working with a colleague as good natured and positive as Debbie. I am also grateful to her for introducing me to the Burning Man Festival, an experience I will never forget!

Many members of the *BABAR* Collaboration have provided sagely advice as I have developed this analysis. In particular I thank Bruce Schumm for his weekly inputs which have been invaluable. I am also grateful to the members of the Radiative Penguin AWG for their regular counsel, especially the conveners John Walsh, Kevin Flood and Jurgen Kroseberg. On the detector operations side I would like to show my appreciation to the members of the Trigger group: Dan Walker for teaching me the workings of the EMT and the system managers Su Dong and Rainer Bartoldus for their advice at operations meetings. Also thanks to Tom Latham for making the EMT a straightforward system to watch over!

California is a fantastic place to live and SLAC a great place to work. My moving there was a smooth process thanks to the UK liaison Jean as well as Robyn and Elaine in the International Office. The Californian experience was of course enhanced massively by the great bunch of people I spent my time with out there, either climbing volcanoes or getting drunk in downtown Palo Alto on 'Swanky Nights'. So a big thanks to Aidan, Christina, Chukwudi, David, Manny, Euginia, Graham, Jen, Jon, Kim, Matt, Mike, Rob, Sudan and Tim. I am beholden to Simon Keenan for his visit, accompanying me to Coachella to witness probably the best festival lineup ever. The trip to Alcatraz was pretty funny too. A particular shout out also goes to Will Panduro Vazquez who not only introduced me to the bay area and its fine eateries, but who has also always been available on short notice for a swift half in London to mutually lament over UK physics funding.

I am particularly grateful to Julia Sedgbeer and Gavin Davies for giving me the opportunity to study with the Imperial HEP group and to Peter Dornan for assigning me to such a great experiment. It's also important to mention my fellow Imperial students, Andy, Chris, the two Davids, Kostas, Joe, Mark and Theo.

While finishing this thesis I have had the pleasure of living with Ali and Matin at 'The Keep' and had the support of many good friends. Thanks to Aidan, Andy, Angelo, Charlotte, Chloe, Dave, Ed (for nodding politely and staring blankly), Eden, Jon, Loz, Matt, Mike, Noemie, Paul, Phil, Rhodri, Rich (for allegedly proof reading), Rosie, Yolanna and anyone else I have forgotton.

A special acknowledgement should go to the UK tax payer (via STFC) for continuing to fund such an important area of research and for paying for my studentship in particular. Despite recent difficulties, the UK is a great place to study particle physics and the interest and support of members of the public in this research is a great motivation for those of us that carry it out.

Finally I would like to thank my family: my mother Joyce, my father Peter and my sister Emma as well as the extended Tibbetts and James families. Claire gets a special mention for joining Mum and Emma in visiting me in California; the road trip to the Grand Canyon will always be one of my fondest memories. I dedicate this thesis to my parents in the hope it partially repays the debt of your time, love and support over the past 28 years.

Declaration

This thesis has been written entirely by the author while in candidature for a research degree at Imperial College. Where arguments and work presented rely on the work of others this has been acknowledged through references to the relevant literature.

With the exception of the coordinate transformation following equation (1.62), which has been calculated by the author, chapter 1 represents the authors best understanding of the relevant physical theory and does not present original work. Where plots have been taken from the existing literature this has been indicated.

The experiment description of chapter 2 represents the authors best understanding of the detector operations environment and does not contain any original work. All plots in this chapter have been provided by the relevant working groups of the *BABAR* collaboration.

The physics analysis presented from chapter 3 onwards has been performed predominantly by the author. It updates an existing physics analysis which is published in [1]. Much of the code used for event reconstruction has been developed by members of the *BABAR* collaboration and implemented by the author based on the previous analysis. Where any part of the analysis relies on the work of others this has been indicated in the text and references to the relevant literature provided.

Work in this thesis uses software tools developed by the *BABAR* collaboration which are not currently documented publicly. Such tools are described in internal *BABAR* Analysis Documents (BADs) which are referenced in this thesis. Furthermore published *BABAR* analyses are generally documented in more detail in support BADs, therefore references to published *BABAR* analyses also contain a support BAD number for completeness. It is likely that the *BABAR* collaboration will make all BADs publically available in the future.

Signed: _____

Date: _____

"Y'know, it's very hard to talk quantum using a language originally designed to tell other monkeys where the ripe fruit is."

Terry Pratchett

Contents

List of Figures

List of Tables

For my family.

Chapter 1

Physics Background

This thesis presents a study of $b \to s/d\gamma$ electroweak radiative penguin decays using e^+e^- collision data from the *BABAR* experiment. Due to quark confinement the process is studied through analysis of the inclusive hadronic decays $B \to X_{s/d}\gamma$, where $X_{s/d}$ represents any charmless hadronic system formed from the resultant s/d quark and spectator quark of the B meson. This chapter motivates the study by discussing the relevant physical theory. An overview of electroweak interactions in the Standard Model (SM) is given and flavour mixing in the quark sector is considered from descriptions of the Cabbibo-Kobayashi-Maskawa (CKM) matrix [2, 3] and resulting Unitarity Triangle (UT) [4]. The phenomenology of inclusive $B \to X_{s/d}\gamma$ decays is discussed through an introduction to rate calculations using operator product expansion. The relation of such rates to the ratio of CKM matrix elements $|V_{td}/V_{ts}|$ then directly motivates the analysis described in subsequent chapters. Finally modelling of the photon energy spectrum is discussed. The following overview of the relevant theory from the SM of particle physics is largely based on the text books [5–7] and the content of lecture courses attended by the author, the proceedings of which are available in [8, 9].

1.1 The Standard Model of Particle Physics

The fundamental constituents of matter, the quarks and leptons, are fermions which form three generations as shown in table 1.1. Their interactions, with the exception of gravity, are described by the Standard Model of particle physics. The SM is constructed by performing gauge transformations on the free Dirac field equation of the fermion spinors and then requiring the transformations be a local space-time symmetry of the Lagrangian. This demands a covariant derivative which, when acting on the fermion

(a)

Q	1	2	3
$-\frac{1}{3}$	d	s	b
$+\frac{2}{3}$	u	c	t

(b)

Q	1	2	3
-1	e^-	μ^-	τ^-
0	ν_e	ν_μ	ν_τ

TABLE 1.1: The fundamental fermions of the SM are (a) the quarks and (b) the leptons. Here they are shown in order of increasing generation from left to right (1-3) and are classified into rows of equal electric charge, Q, given in units of the proton charge.

field, transforms as the field itself. The result is the introduction of terms in the Lagrangian whereby the spinors couple to vector gauge fields. The coupling strengths are arbitrary and there exist as many gauge fields as there are generators of the symmetry group defining the transformation. Kinetic terms for the freely propagating vector fields also appear. The complete SM of strong and electroweak interactions requires local gauge symmetry for transformations under

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y. \tag{1.1}$$

The component $\mathrm{SU}(3)_C$ represents the coupling of the quarks to the eight gluon vector fields through the colour charge. The remaining $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ component forms the Glashow-Salam-Weinberg (GSW) model [10–12] of unified electroweak interactions, the structure of which is described below. Masses of the fermions and weak force mediators are generated in the SM through electroweak spontaneous symmetry breaking (SSB) through the Higgs mechanism [13], requiring the addition of the scalar Higgs field.

To date the SM has been incredibly successful in both accurately predicting and accounting for the phenomenology in all experimental observations of the strong, weak and electromagnetic interactions; however, the SM is not a complete theory, it requires a number of arbitrary constants to be input from experimental measurements as it makes no prediction of their magnitude. Furthermore loop corrections relating to calculations of the Higgs mass are quadratically divergent in the SM[1] unless its value is fine tuned such that the divergences cancel. These are among the motivations for constructing new physical theories which aim to go beyond the Standard Model (BSM) and provide a more complete picture of fundamental physics. All such theories must reduce to the SM in the correct limits so as to be consistent with the experimental data.

[1]The so called hierarchy problem.

1.1.1 Electroweak Interactions

The GSW model of electroweak interactions considers left-handed fermion doublets and right-handed fermion singlets, which for a single generation of quarks and leptons are,

$$\psi_L = \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L ; \; \psi_R = e^-_R, \; u_R, \; d_R; \tag{1.2}$$

plus corresponding antiparticles. Note that there are implicitly no right-handed neutrinos in the SM[2]. Interactions are described by the symmetry group,

$$SU(2)_L \otimes U(1)_Y, \tag{1.3}$$

where $SU(2)_L$ couples the left-handed fermion doublets to the gauge fields of the weak isospin, $\boldsymbol{I}$, and $U(1)_Y$ the fermions to the gauge field of the weak hypercharge, Y. Under (1.3) the spinor fields of (1.2) transform as,

$$\psi_L \rightarrow \psi'_L \equiv e^{iY\beta(x)} U_L(x) \psi_L; \quad U_L(x) \equiv e^{i\frac{\sigma_i}{2}\alpha_i(x)} \quad (i = 1, 2, 3), \tag{1.4}$$

$$\psi_R \rightarrow \psi'_R \equiv e^{iY\beta(x)} \psi_R. \tag{1.5}$$

In (1.4) and (1.5), Y and σ_i (the Pauli spin matrices) are the generators of the groups $U(1)$ and $SU(2)$ respectively. The four gauge parameters of the transformation are $\beta(x)$ representing a local phase change, and $\alpha_i(x)$ representing a local rotation of weak isospin. The required covariant derivatives for the Lagrangian to remain locally invariant are,

$$D^{(L)}_\mu \psi_L \equiv [\partial_\mu + ig\widetilde{W}_\mu(x) + ig' Y B_\mu(x)]\psi_L; \quad \widetilde{W}_\mu(x) = \frac{\sigma_i}{2} W^i_\mu(x), \tag{1.6}$$

$$D^{(R)}_\mu \psi_R \equiv [\partial_\mu + ig' Y B_\mu(x)]\psi_R. \tag{1.7}$$

It can be seen that four vector gauge fields are introduced: the B_μ singlet and the W^i_μ triplet. Their respective coupling constants are g' and g. In order for the covariant derivative to transform as the fermion fields the gauge fields must transform as,

$$B_\mu(x) \rightarrow B'_\mu(x) \equiv B_\mu(x) - \frac{1}{g'}\partial_\mu \beta(x), \tag{1.8}$$

$$\widetilde{W}_\mu(x) \rightarrow \widetilde{W}'_\mu(x) \equiv U_L(x)\widetilde{W}_\mu U^\dagger_L(x) + \frac{i}{g}[\partial_\mu U_L(x)]U^\dagger_L(x). \tag{1.9}$$

[2]Experimental evidence of neutrino flavour mixing suggests neutrinos do have finite mass [14]; however, this is not directly relevant to the discussion here and so ignored.

From the above and by adding the correct gauge invariant kinetic terms for the vector fields the electroweak Lagrangian is therefore,

$$\mathcal{L}_{EW} = i\overline{\psi}_L \gamma^\mu D_\mu^{(L)} \psi_L + i\overline{\psi}_R \gamma^\mu D_\mu^{(R)} \psi_R - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{4} W_{\mu\nu}^i W_i^{\mu\nu}, \tag{1.10}$$

where,

$$B_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu, \tag{1.11}$$

$$W_{\mu\nu}^i \equiv \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g\epsilon^{ijk} W_\mu^j W_\nu^k. \tag{1.12}$$

Here ϵ^{ijk} is the rank three Levi-Civita tensor. Gauge invariance and the different transformation properties of right and left-handed spinors require the fermions and quanta of the gauge fields in (1.10) to be massless. After SSB the gauge fields mix into physical mass eigenstates,

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp i W_\mu^2), \tag{1.13}$$

$$Z_\mu = -B_\mu \sin\theta_W + W_\mu^3 \cos\theta_W, \tag{1.14}$$

$$A_\mu = B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W. \tag{1.15}$$

This introduces the weak mixing angle, θ_W, which is given by $\tan\theta_W = g'/g$. Finally the respective couplings of the weak eigenstates of the fermions to the mass eigenstates of the gauge fields, (1.13), (1.14) and (1.15), can be written down as,

$$-i\frac{g}{\sqrt{2}} \gamma^\mu \frac{1}{2}(1-\gamma^5), \tag{1.16}$$

$$\frac{-ig}{\cos\theta_W} \gamma^\mu \frac{1}{2}[I_3(1-\gamma^5) - 2Q\sin^2\theta_W], \tag{1.17}$$

$$-ieQ\gamma^\mu. \tag{1.18}$$

Here $Q = I_3 + (Y/2)$ is the fermion electric charge. It can be seen that the weak charged current, $W^\pm$, couples only to left-handed fermion fields due to the $(1-\gamma^5)$ term in (1.16). The neutral weak current, Z^0, couples to both left-handed fermions and charged right-handed fermions. Finally the massless photon field, A_μ, only couples to the electric charge.

1.1.2 Charged Current Weak Interactions of Quarks

SSB through the Higgs mechanism allows the introduction of gauge invariant Yukawa terms in the Lagrangian for each fermion,

$$\mathcal{L}_Y = -Y_f \overline{\psi}_L^{(f)} \Phi f_R + h.c., \tag{1.19}$$

where *h.c.* stands for hermitian conjugate, Φ is the scalar Higgs $SU(2)_L$ doublet, Y_f is the Yukawa coupling for fermion f and $\psi_L^{(f)}$ the fermion doublet containing f_L. The Higgs doublet results in mass terms for each fermion (excluding the neutrino which has no right-handed field in the SM),

$$-\sqrt{2}\frac{Y_f M_W}{g}(\overline{f}_L f_R + \overline{f}_R f_L) = \sqrt{2}\frac{Y_f M_W}{g}\overline{f}f, \tag{1.20}$$

where M_W is the mass of the charged current weak propagator. The fermion mass is directly read from (1.20) as,

$$m_f = \sqrt{2}\frac{Y_f M_W}{g} \tag{1.21}$$

The above holds for a single generation; however, in the three generation SM it is possible to write Yukawa terms which mix quarks of different generations. In this scenario (1.19) for the quarks becomes,

$$\mathcal{L}_Y = \sum_{i,j=0}^{3}(\Gamma_{ij}^u \overline{\psi}_L^{(u)i}\Phi u_R^j + \Gamma_{ij}^d \overline{\psi}_L^{(d)i}\Phi d_R^j + h.c.);\ u = \begin{pmatrix} u \\ c \\ t \end{pmatrix},\ d = \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.22}$$

The Γ_{ij}^q matrix elements mix generations i, j and represent couplings analogous to the Y_f of a single generation. It follows that the quark mass matrix has the form,

$$M_{ij}^q = \sqrt{2}\frac{\Gamma_{ij}^q M_W}{g}. \tag{1.23}$$

To obtain the physical mass eigenstates of the quarks the mass matrix is diagonalised through the introduction of unitarity matrices, $V_{L(R)}^q$, such that,

$$V_L^q M^q V_R^{q\dagger} = \mathcal{M}_{diag}^q. \tag{1.24}$$

The relation of the quark mass eigenstates to their weak eigenstates is then given by,

$$u_{L(R)}^m = V_{L(R)}^u u_{L(R)};\quad d_{L(R)}^m = V_{L(R)}^d d_{L(R)}. \tag{1.25}$$

1.1.2.1 The Cabibbo-Kobayashi-Maskawa Matrix

Given (1.25) it is possible to rewrite the electroweak interaction Lagrangian in terms of the quark mass eigenstates. For neutral currents this has no effect, but for charged

currents,

$$\mathcal{L}_{CC} = -\frac{g}{\sqrt{2}}\overline{u}_L\gamma^\mu W^+_\mu d_L + h.c. \tag{1.26}$$

$$= -\frac{g}{\sqrt{2}}\overline{u}^m_L\gamma^\mu W^+_\mu (V^u_L V^{d\dagger}_L) d^m_L + h.c. \tag{1.27}$$

The up-type quarks thus couple to a linear combination of the down-type quarks and the strength of these couplings are determined by the unitary CKM matrix, $(V^u_L V^{d\dagger}_L) = V_{CKM}$ [2, 3]. The CKM matrix can be written as,

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{1.28}$$

From (1.28) it follows that weak charged currents allow transitions between quarks of different generations and that the coupling strengths of these flavour transitions are determined by the weak coupling strength, g, scaled by the relevant CKM matrix element V_{ij}.

The most general form a unitarity 3×3 matrix can take is characterised by nine real parameters: three moduli and six phases; however, in the case of the CKM matrix the $U(1)$ global symmetry of the Lagrangian means five of the phases are unobservable as they can be 'absorbed' by redefining the phases of the quark fields. The CKM matrix is therefore completely described by three moduli and one phase. The standard representation as advocated by the Particle Data Group (PDG) [14] is,

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{-i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix}, \tag{1.29}$$

where $c_{ij} = \cos\theta_{ij},\ s_{ij} = \sin\theta_{ij}$. The angles θ_{ij} are the three moduli and δ_{13} represents the phase; i and j are generation indices. A common representation in the study of B meson decays is that of Wolfenstein [15] using the parameters λ, A, ρ, η whereby,

$$s_{12} \equiv \lambda;\ s_{23} \equiv A\lambda^2;\ s_{13}e^{i\delta_{13}} \equiv A\lambda^3(\rho - i\eta). \tag{1.30}$$

Taking $\lambda \approx 0.22$ as an expansion parameter the form of (1.29) in this representation, neglecting terms $\geq \mathcal{O}(\lambda^4)$ is,

$$V_{CKM} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix}. \tag{1.31}$$

Higher order expansions of the Wolfenstein parameterisation to $\mathcal{O}(\lambda^6)$ exist and under this scheme, to keep concise notation, it is convenient to define the terms [16],

$$\overline{\rho} = \rho\left(1 - \frac{\lambda^2}{2}\right);\ \overline{\eta} = \eta\left(1 - \frac{\lambda^2}{2}\right). \tag{1.32}$$

1.1.2.2 *CP* Violation

Any local quantum field theory constructed in the Lagrangian formalism can be subject to three discrete transformations: parity (P), time reversal (T) and charge conjugation (C). Parity performs the operation $P : (t \to t;\ \mathbf{x} \to -\mathbf{x})$ and time reversal the operation $T : (t \to -t;\ \mathbf{x} \to \mathbf{x})$. Charge conjugation substitutes particle for antiparticle, reversing the sign on all internal quantum numbers of the particle while leaving the four-momentum, space-time coordinate and spin unchanged. The combined operation CPT is always a symmetry of the Lorentz invariant Lagrangian [17].

It is possible to rewrite (1.27) as

$$\mathcal{L} = -\frac{g}{\sqrt{2}}[\overline{u}_i\gamma^\mu W_\mu^+(1-\gamma^5)V_{ij}d_j + \overline{d}_j\gamma^\mu W_\mu^-(1-\gamma^5)V_{ij}^* u_i]. \tag{1.33}$$

Under the discrete transformation CP the terms in this equation transform as

$$\overline{\psi}_i\gamma^\mu W_\mu(1-\gamma^5)\psi_j \xrightarrow{CP} \overline{\psi}_j\gamma^\mu W_\mu(1-\gamma^5)\psi_i; \tag{1.34}$$

however, V_{ij} and V_{ij}^* do not transform. It directly follows that the remaining phase in the CKM matrix can result in the combined CP transformation not being a symmetry of the Lagrangian. It is also required that none of the quarks are mass degenerate and that none of the three CKM moduli are zero or $\pi/2$ as either would decouple the flavour transition between at least two of the quarks allowing the final CKM phase to be absorbed by a redefinition of quark field phase. CP asymmetry was first observed in 1964 in the neutral kaon system [18] and has since been seen by the *BABAR* experiment in the neutral B meson system [19, 20]. Furthermore the *BABAR* collaboration have also claimed evidence of CP asymmetry in the charged B meson system [21]. All these results are consistent with the CKM formalism of the SM.

Experimentally there are three types of observable CP violation in the SM arising from the CKM phase factor [4]. CP violation in decay, known as direct CP violation, occurs when the amplitude, A, of a decay is not identical to the conjugate decay amplitude $\overline{A}$. CP violation in mixing, known as indirect CP violation, occurs in the mixing of neutral mesons where the flavour eigenstates are different from the mass eigenstates. This can be observed through the neutral meson weak eigenstate and its anti-particle

partner decaying to the same CP eigenstate with different amplitudes. The decay $B^0 \to J/\psi K_S$ and its charge conjugate are an example of such a system. The final type of CP violation occurs from a quantum mechanical interference of the first two. Theoretically CP violation can occur in $B \to X_d\gamma$ transitions [22]; however, the *BABAR* dataset is insufficient to render the required statistical sensitivity for such a measurement.

1.1.2.3 The Unitarity Triangle

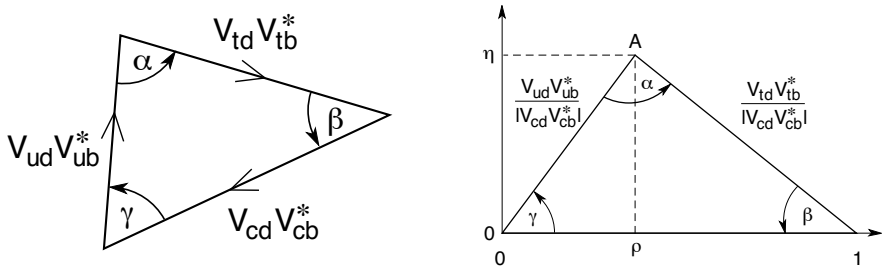


FIGURE 1.1: Sketch of the Unitarity Triangle from [4] showing it both as (a) an arbitrary triangle in the complex plane and (b) normalised such that the apex lies at (ρ, η).

CP violation in the SM is constrained by the unitarity of the CKM matrix, $V_{CKM}V_{CKM}^\dagger = I$. It follows that testing the constraints of unitarity can test the SM picture of CP violation. There exist six orthogonality relations from unitarity which are each expressed as the vanishing sum of three complex numbers and hence represented as six closed triangles in the complex plane. Of particular interest to B meson decays is the relation,

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{1.35}$$

In the complex plane this relation is represented by the so called Unitarity Triangle which is shown in figure 1.1 [4]. It can be seen that it is convenient to normalise the UT by dividing each side by a factor $V_{cd}V_{cb}^*$ which results in the apex having coordinate (ρ, η) in the Wolfenstein representation. This holds for higher order too so one could easily write $(\overline{\rho}, \overline{\eta})$ as the apex. For very precise experimental measurements of CKM matrix elements and UT triangle geometry the higher order parameterisation is preferred due to its increased analytical accuracy.

Experimentally it is possible to independently measure the angles α, γ and β as well as the triangle sides. Performing such measurements allows one to determine whether they are consistent with a closed triangle and therefore whether the CKM matrix is unitary and fully accounts for experimental observations of CP violation. The current

Param	Value
λ	$0.2257^{+0.0009}_{-0.0010}$
A	$0.814^{+0.021}_{-0.022}$
$\overline{\rho}$	$0.135^{+0.031}_{-0.016}$
$\overline{\eta}$	$0.349^{+0.015}_{-0.017}$

TABLE 1.2: Fitted values of the Wolfenstein parameters from experimental constraints with associated errors [14].

published experimental data are consistent with the UT closing. Measurements for the Wolfenstein parameters are given in table 1.2 [14]. The analysis in this thesis is used to extract a measurement of the ratio of CKM parameters $|V_{td}/V_{ts}|$, the details of which are discussed below. It can be seen from (1.31) that in the Wolfenstein parameterisation,

$$\left|\frac{V_{td}}{V_{ts}}\right| = \lambda\sqrt{(1-\rho)^2+\eta^2}. \tag{1.36}$$

Hence $|V_{td}/V_{ts}|$ is directly proportional to the length of the side of the UT between (ρ,η) and $(1,0)$ with λ as the scaling factor. The form of (1.36) at higher order with $\rho \to \overline{\rho}$ and $\eta \to \overline{\eta}$ is discussed below. Typically (1.36) is constrained experimentally from $|V_{td}/V_{ts}| \propto (\Delta m_d/\Delta m_s)$ where $\Delta m_{d(s)}$ is the mixing frequency of the neutral $B^0_{d(s)}\overline{B}^0_{d(s)}$ meson system [14]. An independent constraint can be obtained through measuring the rates of electroweak penguin decays and this is the aim of the physics analysis presented in subsequent chapters. Current experimental constraints on $|V_{td}/V_{ts}|$ from these different types of analysis are discussed in section 1.3.

1.2 Phenomenology of $B \to X_{s/d}\gamma$ Decays

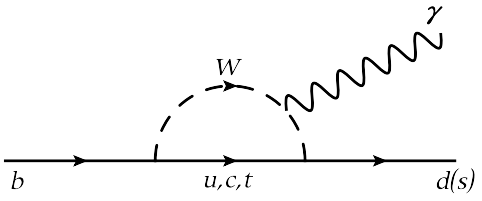


FIGURE 1.2: Feynman diagram showing the SM loop process dominating the FCNC transition $b \to s/d\gamma$.

The electroweak formalism of the SM results in there being no flavour changing neutral currents (FCNC) at tree level. The transition $b \to s/d\gamma$ proceeds at leading order through one-loop penguin processes involving the emission and reabsorption of a W boson combined with an intermediate up-type quark as shown in figure 1.2. The CKM

and mass dependance of the quark propagator in the loop results in different contributions from each of u, c and t, with the more massive t quark dominating. The process $b \to s/d\gamma$ is thus higher order and suppressed in the SM, with the transition $b \to d\gamma$ itself suppressed relative to its CKM favoured counterpart, $b \to s\gamma$. It is possible that contributions from BSM physics can lead to a measurable difference between the experimental rate and the calculated rate from pure SM contributions. Figure 1.3 shows

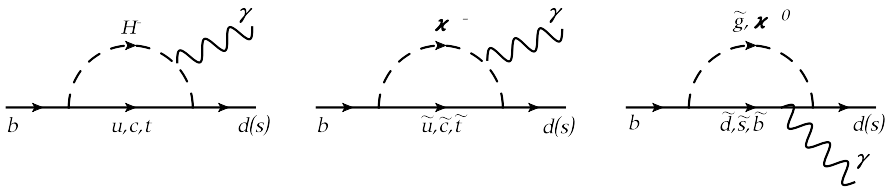


FIGURE 1.3: Examples of possible BSM contributions to $b \to s/d\gamma$. The photon can radiate from any charged leg.

possible contributions to the loop from certain supersymmetric models. The presence of such particles can lead to a departure from unitarity of the flavour structure in the quark sector, hence the measurement of these processes perform a stringent test on the validity of the SM at higher order.

The pure electroweak process above is studied through the weak decays of B mesons of the form $B \to X_{s/d}\gamma$ due to quark confinement. Consequently calculations of the expected SM rates must account for QCD corrections due to strong interactions between the quarks and gluons. Such corrections are complicated by the non-perturbative nature of long range QCD effects which are important when considering hadron decays. The calculation is therefore a non-trivial combination of the electroweak and strong interactions. In order to solve this the formalism of operator product expansion (OPE) is employed.

1.2.1 Operator Product Expansion

Consider the amplitude A for a general weak meson i decaying to final state f. In the OPE formalism this amplitude is expressed as [23],

$$A(i \to f) = \langle f | \mathcal{H}_{eff} | i \rangle = \frac{G_F}{\sqrt{2}} \sum_i V^i_{CKM} C_i(\mu, M_W) \langle f | Q_i(\mu) | i \rangle , \tag{1.37}$$

where $G_F = (\sqrt{2}g^2/8M_W^2)$ is the Fermi constant. The decay amplitude is modelled through an effective theory governed by the Hamiltonian $\mathcal{H}_{eff}$. This factorises the problem into Wilson coefficient functions, C_i and the matrix elements of local operators Q_i at a factorisation scale μ. Physically the Wilson coefficients contain the information

on physics over short distances such as heavy virtual particle exchange (e.g. the W boson or top quark) and short range QCD effects which due to asymptotic freedom can be treated perturbatively. The Q_i operators describe the long range interactions, in particular non-perturbative QCD effects which are important in hadronic decays. By definition A is independent of μ hence the scale dependance of $C_i(\mu)$ must cancel that of $Q_i(\mu)$. For B meson decays it is usual to take $\mu \sim \mathcal{O}(m_b)$ where m_b is the mass of the b quark.

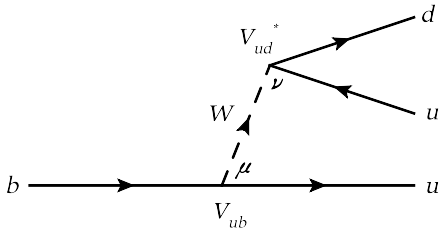


FIGURE 1.4: Feynman diagram showing the tree level weak transition $b \to ud\overline{u}$.

To illustrate the construction of such an effective theory consider the simple case of the quark level transition $b \to ud\overline{u}$ shown in figure 1.4 which can contribute to the combinatoric backgrounds of $B \to X_{s/d}\gamma$ decays if there is a π^0 or η meson in the final state. The evaluation of this Feynman diagram yields,

$$-\frac{g^2}{8}V_{ud}^*V_{ub}[\overline{d}\gamma^\nu(1-\gamma^5)u]\left[\frac{g_{\nu\mu}}{k^2-M_W^2}\right][\overline{u}\gamma^\mu(1-\gamma^5)b], \tag{1.38}$$

where k is the four-momentum transfer due to the weak boson. Since $k^2 < m_b^2 \ll M_W^2$ the following approximation is made,

$$\frac{g_{\nu\mu}}{k^2-M_W^2} \longrightarrow -\frac{g_{\nu\mu}}{M_W^2} \equiv -\left(\frac{8G_F}{\sqrt{2}g^2}\right)g_{\nu\mu}. \tag{1.39}$$

Thus the effective Hamiltonian can be written as,

$$\mathcal{H}_{eff} = \frac{G_F}{\sqrt{2}}V_{ud}^*V_{ub}(\overline{d}_\alpha u_\alpha)_{V-A}(\overline{u}_\beta b_\beta)_{V-A} \equiv \frac{G_F}{\sqrt{2}}V_{ud}^*V_{ub}O_1, \tag{1.40}$$

where α and β are colour indices and the $V-A$ subscript implies the omitted $\gamma^\mu(1-\gamma^5)$ factors. Essentially the W boson has been removed as a degree of freedom from the problem and the transition is described by the local four quark current-current operator O_1, shown in figure 1.5.

In this trivial case without QCD corrections the corresponding Wilson coefficient is unity. To account for QCD corrections the effective theory is extended by considering gluon exchanges, examples of which are shown in figure 1.6. The factorisable corrections in

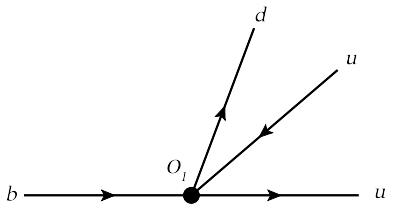


FIGURE 1.5: Four quark current-current operator for the weak transition $b \to ud\overline{u}$.

figure 1.6(a) result in the Wilson coefficient C_1 acquiring scale dependance, i.e. $C_1(\mu) \neq 1$. The non-factorisable corrections of figure 1.6(b) require the addition of a second current-current operator from operator mixing,

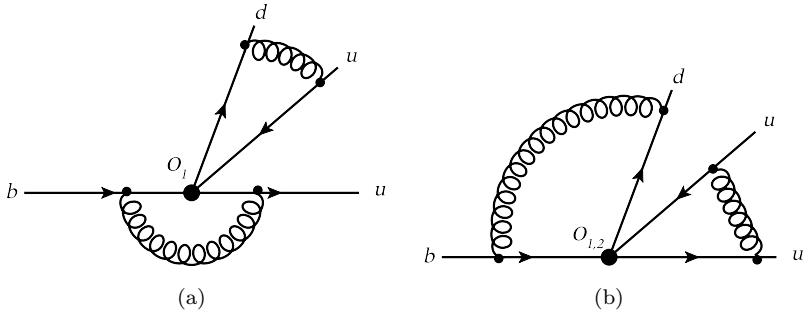


FIGURE 1.6: QCD corrections to $b \to ud\overline{u}$ showing (a) factorisable and (b) non-factorisable contributions.

$$O_2 \equiv (\overline{u}_\alpha d_\beta)_{V-A}(\overline{u}_\beta b_\alpha)_{V-A}. \tag{1.41}$$

Hence the effective Hamiltonian becomes,

$$\mathcal{H}_{eff} = \frac{G_F}{\sqrt{2}} V_{ud}^* V_{ub}[C_1(\mu)O_1 + C_2(\mu)O_2]. \tag{1.42}$$

Once the effective Hamiltonian has been constructed, as in the above example, the problem becomes one of evaluating the Wilson coefficients and local operator matrix elements. For the latter there are eight relevant operators in the evaluation of $B \to X_s\gamma$ [24]. Firstly the two current-current operators Q_1 and Q_2,

$$Q_1 = (\overline{s}T^a c)_{V-A}(\overline{c}T_a b)_{V-A}, \tag{1.43}$$

$$Q_2 = (\overline{s}c)_{V-A}(\overline{c}b)_{V-A}, \tag{1.44}$$

where T_a $(a = 1 \ldots 8)$ are the $SU_c(3)$ generators and imply the colour indices of the analogous operator O_2 above. Secondly there are four QCD penguin operators which arise from diagrams such as that shown in figure 1.7 and have the form,

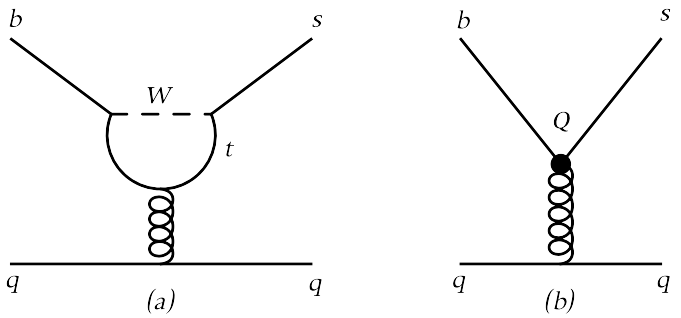


FIGURE 1.7: The QCD penguin contributions to $B \to X_s\gamma$, (a) the full SM contribution and (b) the effective operator Q with the heavy degrees of freedom removed. Only the dominant top quark loop is shown.

$$Q_3 = (\overline{s}b)_{V-A} \sum_q (\overline{q}q)_{V-A}, \tag{1.45}$$

$$Q_4 = (\overline{s}T^a b)_{V-A} \sum_q (\overline{q}T_a q)_{V-A}, \tag{1.46}$$

$$Q_5 = (\overline{s}b)_{V-A} \sum_q (\overline{q}q)_{V+A}, \tag{1.47}$$

$$Q_6 = (\overline{s}T^a b)_{V-A} \sum_q (\overline{q}T_a q)_{V+A}. \tag{1.48}$$

These differ from the current-current operators as there are additional right handed contributions from $\gamma^\mu(1+\gamma^5)$ terms (denoted $V+A$) and there is a sum over all $q\overline{q}$ pairs that the gluon can produce. Finally there are contributions from photon and gluonic magnetic penguin operators shown in figure 1.8 and which have the form,

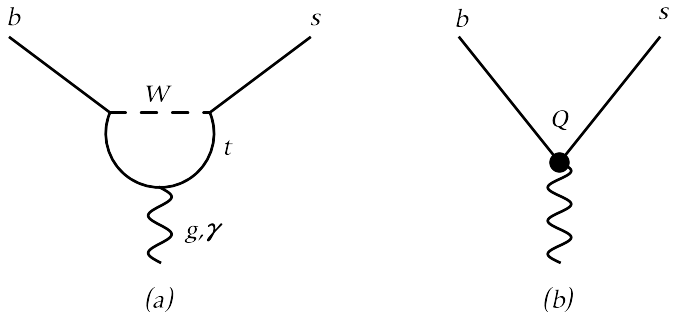


FIGURE 1.8: The magnetic penguin contributions to $B \to X_s\gamma$, (a) the full SM contribution and (b) the effective operator Q with the heavy degrees of freedom removed. The gauge boson leg is a photon for Q_7 and a gluon for Q_8.

$$Q_7 = \frac{e}{8\pi^2} m_b(\mu) \overline{s}\sigma^{\mu\nu}(1+\gamma^5) b F_{\mu\nu}, \tag{1.49}$$

$$Q_8 = \frac{g_s}{8\pi^2} m_b(\mu) \overline{s}\sigma^{\mu\nu} T^a (1+\gamma^5) b G^a_{\mu\nu}, \tag{1.50}$$

where e and $F_{\mu\nu}$ (g_s and $G^a_{\mu\nu}$) are the electromagnetic (strong) coupling constant and field strength tensor, respectively. For Q_7 and Q_8, $m_b(\mu)$ is the running b quark mass. The equivalent Q_i operators for $B \to X_d\gamma$ are trivially obtained by substituting the s quark above for a d quark.

1.2.1.1 Heavy Quark Expansion

The Q_i operators represent the non-perturbative contributions to the decay $B \to X_{s/d}\gamma$. In the calculation of exclusive B meson decays the evaluation of Q_i can represent a significant and complex challenge; however, in the case of the inclusive decays considered here the problem becomes much simpler. The branching ratio can be calculated using an expansion in inverse powers of m_b [25],

$$\mathcal{B}(B \to X_q\gamma) = \mathcal{B}(b \to q\gamma) + \mathcal{O}\left(\frac{1}{m_b^2}\right). \tag{1.51}$$

This is known as heavy quark expansion (HQE). The leading term in (1.51) represents the b quark decay which can be calculated in perturbation theory. In the limit $m_b \to \infty$ the equivalence to the quark level decay is exact and calculations show that the $\mathcal{O}\left(\frac{1}{m_b^2}\right)$ corrections are $\sim 3\%$ [22]. Essentially the approximation relies on the fact that the b quark decay is a short distance process whose decay time is much less than the timescale of hadronisation to the final state X_q. The advantage of the HQE approximation is that the Q_i operators can effectively be calculated perturbatively. In particular their scale dependence can be calculated and the cancellation of this dependance with the corresponding $C_i(\mu)$ investigated. It follows that under this scheme the long range QCD contributions to the decay $B \to X_{s/d}\gamma$ are well under control making the decay a laboratory for the short range electroweak effects contained within the Wilson coefficients.

1.2.1.2 Wilson Coefficient Evaluation

The $C_i(\mu)$, which contain the short range QCD effects, are always calculable in perturbation theory due to asymptotic freedom of the strong coupling $\alpha_s(\mu)$, which at scales $\mu \geq \mathcal{O}(1\,\text{GeV})$ can be used as an expansion parameter; however, at the scale $\mathcal{O}(m_b) \ll M$ ($M = M_W, m_t$) large logarithms of the form $\ln(M/\mu)$ multiply α_s due to the consequences of evaluating a renormalisable quantum field theory when different scales are present [17]. To maintain the validity of perturbation theory, a renormalisation group analysis is performed allowing summation of the logarithm terms to all orders [25]. The $C_i(\mu)$ are therefore evaluated with renormalisation group improved perturbation theory. The leading order (LO) term in such an analysis sums $[\alpha_s \ln(M/\mu)]^n$

contributions. LO calculations can contain large theoretical uncertainties due to leftover μ dependance in the amplitude calculation. This is much reduced by additionally summing $\alpha_s \left[\alpha_s \ln(M/\mu)\right]^n$ contributions giving a next to leading order (NLO) calculation. NLO evaluations for all the $C_i(\mu)$ relevant to the decay $B \to X_{s/d}\gamma$ have been made [24].

1.2.2 Calculation of $|V_{td}/V_{ts}|$

OPE provides a framework to calculate expected SM rates for $B \to X_{s/d}\gamma$ and the corresponding charge conjugate decay, $\overline{B} \to \overline{X}_{s/d}\gamma$. The average branching ratio is defined as,

$$\langle \mathcal{B}(B \to X_{s/d}\gamma) \rangle = \frac{\mathcal{B}(B \to X_{s/d}\gamma) + \mathcal{B}(\overline{B} \to \overline{X}_{s/d}\gamma)}{2}. \tag{1.52}$$

It is possible to relate the ratio,

$$R(d\gamma/s\gamma) \equiv \frac{\langle \mathcal{B}(B \to X_d\gamma) \rangle}{\langle \mathcal{B}(B \to X_s\gamma) \rangle}, \tag{1.53}$$

to the CKM parameter $|V_{td}/V_{ts}|$ as shown by Ali et al [22]. A brief summary of their derivation of this relation in terms of the Wolfenstein parameters $\overline{\rho}$ and $\overline{\eta}$ follows.

1.2.2.1 Decay Rates in $B \to X\gamma$

The effective Hamiltonian for the decay $\overline{B} \to \overline{X}_s\gamma$ leads to the following expression for the decay rate,

$$\Gamma(b \to s\gamma) = \frac{m_b^5 G_F^2 \alpha_{EM}}{32\pi^4} |D(b \to s\gamma)|^2, \tag{1.54}$$

$$D(b \to s\gamma) = \lambda_t (A^t_{Re} + iA^t_{Im}) + \lambda_u (A^u_{Re} + iA^u_{Im}), \tag{1.55}$$

where $\lambda_q \equiv V_{qb}V^*_{qs}$ are the relevant CKM factors and the real functions A^q are constructed from the Wilson coefficients and matrix elements. The branching ratio is expressed in terms of the measured semileptonic branching ratio $\mathcal{B}(B \to Xl\nu_l)$,

$$\mathcal{B}(\overline{B} \to \overline{X}_s\gamma) = \frac{\Gamma(\overline{B} \to \overline{X}_s\gamma)}{\Gamma_{sl}} \mathcal{B}(B \to Xl\nu_l), \tag{1.56}$$

where Γ_{sl} is the calculated semileptonic rate. This is now written to explicitly show dependancies on the CKM matrix factors,

$$\mathcal{B}(\overline{B} \to \overline{X}_s\gamma) = \frac{1}{|V_{cb}|^2} [D_t|\lambda_t|^2 + D_u|\lambda_u|^2 + D_r Re(\lambda^*_t\lambda_u) + D_i Im(\lambda^*_t\lambda_u)], \tag{1.57}$$

where the D factors, which depend on various input parameters such as m_t, m_b, m_c, μ and α_s, are calculated numerically in table 1 of [22]. The average branching ratio is obtained from dropping the last term of (1.57) due to the corresponding conjugate term having opposite sign. In fact due to the dominance of the leading term it is possible to write,

$$\langle \mathcal{B}(B \to X_s\gamma) \rangle = \frac{|\lambda_t|^2}{|V_{cb}|^2} D_t. \tag{1.58}$$

Following the same procedure for $B \to X_d\gamma$ gives,

$$\langle \mathcal{B}(B \to X_d\gamma) \rangle = \frac{1}{|V_{cb}|^2} [D_t|\xi_t|^2 + D_u|\xi_u|^2 + D_r Re(\xi_t^* \xi_u)], \tag{1.59}$$

where $\xi_q \equiv V_{qb}V_{qd}^*$. The D factors of (1.59) are identical to those of (1.57), however, the CKM factors of (1.59) reduce the dominance of the leading term so all terms must be retained.

1.2.2.2 $|V_{td}/V_{ts}|$ from Inclusive Decays

From (1.58) and (1.59) it follows that,

$$R(d\gamma/s\gamma) = \frac{|\xi_t|^2}{|\lambda_t|^2} + \frac{D_u}{D_t}\frac{|\xi_u|^2}{|\lambda_t|^2} + \frac{D_r}{D_t}\frac{Re(\xi_t^*\xi_u)}{|\lambda_t|^2}. \tag{1.60}$$

This is rewritten in terms of the Wolfenstein parameters $\overline{\rho}$ and $\overline{\eta}$ [22],

$$R = \lambda^2[1 + \lambda^2(1 - 2\overline{\rho})] \left[(1-\overline{\rho})^2 + \overline{\eta}^2 + \frac{D_u}{D_t}(\overline{\rho}^2 + \overline{\eta}^2) + \frac{D_r}{D_t}(\overline{\rho}(1-\overline{\rho}) - \overline{\eta}^2) \right]. \tag{1.61}$$

Comparing the leading term of (1.60) to (1.61), the expression for $|V_{td}/V_{ts}|$ as a function of the higher order Wolfenstein parameters is found to be,

$$\frac{|\xi_t|^2}{|\lambda_t|^2} = X^2 = \lambda^2 \left[1 + \lambda^2(1 - 2\overline{\rho})\right] \left[(1-\overline{\rho})^2 + \overline{\eta}^2\right], \tag{1.62}$$

where the definition $X = |V_{td}/V_{ts}|$ is used to simplify the notation. Remaining terms in (1.61) are also functions of $\overline{\rho}$ and $\overline{\eta}$. If one were to measure R and calculate X while constraining these terms with the current Wolfenstein parameter measurements then the value of X obtained would not be truly independent. This is due to previous measurements of X being used to evaluate the world averages of $\overline{\rho}$ and $\overline{\eta}$. For this analysis the extracted value of X is required to be independent. It is therefore proposed to rewrite (1.61) as a function of X and an orthogonal coordinate which is chosen to be the UT angle β. It is noticed from (1.62) that in the higher order Wolfenstein expansion X is no longer directly proportional to the relevant side of the UT as was the case for (1.36). Therefore redefinition of the coordinate system requires substitutions of the

form,

$$(1-\overline{\rho}) = \frac{X}{\lambda}\cos\beta + \lambda^2 f_{\overline{\rho}}(X,\beta), \tag{1.63}$$

$$\overline{\eta} = \frac{X}{\lambda}\sin\beta + \lambda^2 f_{\overline{\eta}}(X,\beta), \tag{1.64}$$

where the f functions account for the higher order corrections to (1.36). From (1.62), the relation $\tan\beta = \overline{\eta}/(1-\overline{\rho})$ which is obtained from the UT geometry, assuming $(X/\lambda) \sim 1$ and neglecting terms with multiplying factors $> \mathcal{O}(\lambda^4)$, the f functions are derived to be,

$$f_{\overline{\rho}}(X,\beta) = \frac{1}{2}\frac{X}{\lambda}\cos\beta\left[1 - 2\frac{X}{\lambda}\cos\beta\right], \tag{1.65}$$

$$f_{\overline{\eta}}(X,\beta) = \frac{1}{2}\frac{X}{\lambda}\sin\beta\left[1 - 2\frac{X}{\lambda}\cos\beta\right]. \tag{1.66}$$

Substituting for $(1-\overline{\rho})$ and $\overline{\eta}$ in (1.61), again assuming $(X/\lambda) \sim 1$ and neglecting terms with multiplying factors $> \mathcal{O}(\lambda^4)$, R is found to be a quadratic in X,

$$R = \kappa_1 X^2 + \kappa_2 X + \kappa_3, \tag{1.67}$$

where,

$$\kappa_1 = 1 + \frac{D_u}{D_t}\left(1 - 2\lambda^2\cos^2\beta\right) - \frac{D_r}{D_t}\left(\lambda^2\cos^2\beta + 1\right), \tag{1.68}$$

$$\kappa_2 = \lambda\cos\beta\left[\frac{D_u}{D_t}\left(3\lambda^2 - 2\right) + \frac{D_r}{D_t}\left(1 + \frac{\lambda^2}{2}\right)\right], \tag{1.69}$$

$$\kappa_3 = \lambda^2\frac{D_u}{D_t}\left(1-\lambda^2\right). \tag{1.70}$$

Numerical calculation of the above κ factors uses the world average measurements of λ and $\sin 2\beta$ [14]. The factors D_u/D_t and D_r/D_t are taken from table 1 of [22] assuming $\mu = 2.5\,\text{GeV}$. The current world average of $m_c/m_b = 0.30^{+0.02}_{-0.03}$ does not constrain the numerical variations of the D factors as a function of m_c/m_b quoted in [22] and hence uncertainty estimates for each D factor are made based on this variation. The complete set of input parameters and their associated uncertainties are show in table 1.3 and give $\kappa_1 = 1.4144$, $\kappa_2 = -0.1049$ and $\kappa_3 = 0.0048$. Given the numerical evaluation of the κ factors, (1.67) is inverted to obtain X as a function of R. The propagation of uncertainties due to the input variables are estimated for the inverted equation. For each input parameter, Z, the other parameters are held constant at their central values and X is recalculated using $Z \pm 1\sigma_Z$. Thus two extreme values of X are obtained and the average of the difference of these extremes relative to the central value of X is taken as the error on X due to the variation in Z. The parameters D_u/D_t and D_r/D_t are

Param	Value	Error
λ	0.2257	$^{+0.0009}_{-0.0010}$
$\sin 2\beta$	0.681	± 0.025
$\cos\beta$	0.931	$^{+0.006}_{-0.007}$
$\frac{D_u}{D_t}$	0.0995	$^{+0.0005}_{-0.0002}$
$\frac{D_r}{D_t}$	-0.310	$^{+0.090}_{-0.062}$

TABLE 1.3: Values of input parameters for the evaluation of the numerical κ factors required to calculate X.

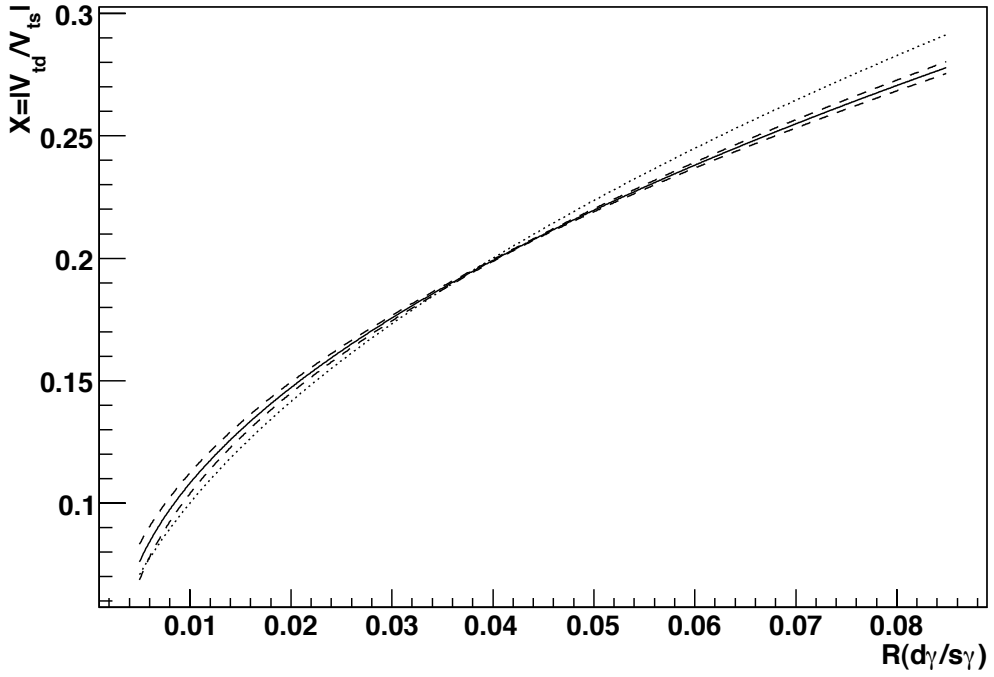


FIGURE 1.9: Calculated variation of $|V_{td}/V_{ts}|$ as a function of $R(d\gamma/s\gamma)$. The solid line is the expected value with the dashed lines showing calculated error. The dotted line shows $|V_{td}/V_{ts}| = \sqrt{R}$

varied simultaneously as their dependence on the uncertainty in m_c/m_b results in their uncertainties being 100% correlated. The three individual uncertainties are added in quadrature to obtain an overall error on X. The numerical evaluation of (1.67) with corresponding uncertainties is shown in figure 1.9. For the range of R plotted, the relative uncertainty on X is seen to vary with a maximum of $\mathcal{O}(9\%)$ at lower R and a minimum of $\mathcal{O}(0.2\%)$ towards the centre of the distribution. Also plotted is the leading term of (1.60), $X = \sqrt{R}$ to illustrate the effect of higher order corrections. The dominant uncertainties in figure 1.9 are due to uncertainties in the numerical evaluation of the D factors. It follows that when corrections from the D factors become less important,

i.e. when the calculated distribution coincides with the leading order distribution at $R \approx 0.04$, the theoretical uncertainty is significantly reduced.

1.2.2.3 $|V_{td}/V_{ts}|$ from Exclusive Resonant $B \to X_{s/d}\gamma$ Decays

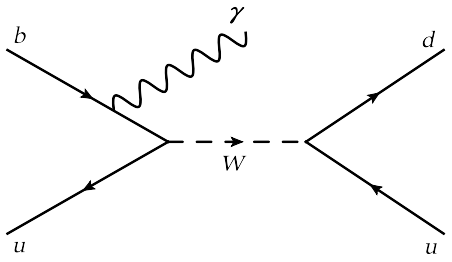


FIGURE 1.10: Feynman diagram showing weak annihilation contributions in the SM to the decay $B^- \to \rho^-\gamma$.

Previous measurements have used radiative penguin decays to extract X by taking the ratio of exclusive resonant decays and applying the calculated theoretical relation [26],

$$\frac{\Gamma(B \to \rho\gamma)}{\Gamma(B \to K^*\gamma)} = S_\rho \left|\frac{V_{td}}{V_{ts}}\right|^2 \left(\frac{1 - m_\rho^2/M_B^2}{1 - m_{K^*}^2/M_B^2}\right)^3 \zeta^2[1 + \Delta R]. \tag{1.71}$$

Here S_ρ is an isospin factor dependent on the charge of the ρ meson, $m_{\rho(K^*)}$ is the mass of the $\rho(K^*)$ meson, M_B the mass of the B meson, ζ a form factor ratio, and ΔR a correction factor for weak annihilation contributions which are shown in figure 1.10. To date such measurements show no discrepancy between the extracted values of X compared to those measured from B meson mixing. However, they are not as competitive as the latter due to limited statistics in the measurement of $B \to \rho\gamma$ and due to a relatively large theoretical error in ζ of $\sim 8\%$. The independence of different operators contributing to the relevant rate calculations for these respective processes means BSM physics could be observable in only one of these measurements and hence the increased uncertainty in radiative penguin measurements does not diminish the importance of their measurement. The significantly reduced theoretical error from the ratio of inclusive decays in extracting X means that the value obtained from the analysis presented here will be competitive with the measurement from exclusive decays.

1.2.3 Kagan-Neubert Model of Photon Energy Spectrum

The experimental interpretation of inclusive $B \to X_{s/d}\gamma$ decays must take account of the spectrum of photon energy[3], E_γ. A number of theoretical schemes can be used to

[3]Defined in the B meson rest frame.

estimate this spectrum for $B \to X_s\gamma$ decays [27–29]. This analysis uses the model of Kagan and Neubert (KN) which is comprehensively described in [27] and summarised here. Although older than other schemes it is chosen as it is readily implemented in the *BABAR* simulation framework and has been shown to be in reasonable agreement with previous *BABAR* experimental investigations [30].

1.2.3.1 Photon Energy Spectrum

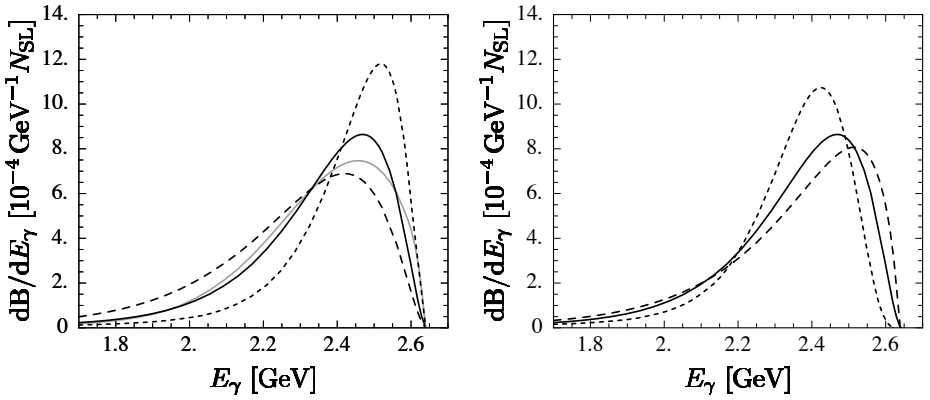


FIGURE 1.11: Plots taken from Figure 3 of [27] showing normalised photon spectra for $B \to X_s\gamma$ decays using various numerical parameterisations of the shape function $F(k)$ defined in (1.74). The left plot shows variations in m_b with $4.65\,\mathrm{GeV}/c^2$ the long-dashed line, $4.8\,\mathrm{GeV}/c^2$ the solid line and $4.95\,\mathrm{GeV}/c^2$ the short-dashed line. The grey curve represents a Gaussian parameterisation of $F(k)$ and is not relevant to the discussion here. The right plot shows variations in μ_π^2 with $0.15\,\mathrm{GeV}^2$ the short-dashed line, $0.30\,\mathrm{GeV}^2$ is the solid line and $0.45\,\mathrm{GeV}^2$ the long-dashed line.

The KN model calculates $\mathcal{B}(B \to X_s\gamma)$ at NLO and then considers the phenomenological contribution of so called 'Fermi motion' of the b quark to HQE. This motion is modelled as a shape function, $F(k)$, which governs the light-cone momentum distribution, k, of the b quark within the B meson which has mass m_B. The photon spectrum is then calculated as a function of the parameter $y = 2E_\gamma/m_B$ from a convolution of the parton model spectrum $P_p(y_p)$ and the shape function,

$$P(y)\,\mathrm{d}y = \int \mathrm{d}k\, F(k)[P_p(y_p)\,\mathrm{d}y_p]_{y_p=y_p(k)}. \tag{1.72}$$

Here $y_p(k) = 2E_\gamma/m_B^*$ with effective mass $m_B^* = m_B + k$ ensuring the Fermi motion model gives the true kinematic limit $E_\gamma^{max} = m_B/2$ rather than $E_\gamma^{max} = m_b/2$ expected from the parton model.

The shape function is found to be theoretically constrained by the moments,

$$A_n = \int \mathrm{d}k\, k^n F(k), \tag{1.73}$$

with $A_0 = 1$, $A_1 = 0$ and $A_2 = \frac{1}{3}\mu_\pi^2$. The KN model choses a parameterisation of $F(k)$ consistent with these constraints,

$$F(k) = N(1-x)^a e^{(1+a)x}; \quad x = \frac{k}{\Lambda}, \tag{1.74}$$

where N is a normalisation factor fixed by A_0, $\Lambda = m_B - m_b$ and $A_2 = \Lambda^2/(1+a)$. It follows that $F(k)$ is parameterised by (m_b, μ_π^2).

Figure 1.11 shows various normalised evaluations of the photon spectra for $B \to X_s\gamma$ decays. The effects of varying each shape function parameter, (m_b, μ_π^2), while keeping the other constant are shown. This analysis models the photon spectrum with ($4.65\,\mathrm{GeV}/c^2$,$0.52\,\mathrm{GeV}^2$) based on fits to $b \to s\gamma$ and $b \to cl\nu$ data [31]. For systematic studies these parameters are varied in a correlated way[4] , according to their measured uncertainties, to give alternative photon spectra with ($4.60\,\mathrm{GeV}/c^2$,$0.60\,\mathrm{GeV}^2$) and ($4.70\,\mathrm{GeV}/c^2$,$0.45\,\mathrm{GeV}^2$).

1.2.3.2 Hadronic Mass Spectrum

When considering different frames of reference (such as the B meson rest frame, the $\Upsilon(4S)$ rest frame and the laboratory rest frame) it is useful to parameterise the photon energy spectrum as a hadronic mass spectrum of the $X_{s/d}$ system, m_X. This is because the Lorentz invariance of m_X results in identical spectra in all rest frames. In the B meson rest frame,

$$m_X^2 = m_B^2 - 2m_B E_\gamma. \tag{1.75}$$

The left hand plot of figure 1.12 shows the mass spectrum corresponding to each of the photon spectra in the left hand plot of figure 1.11. The treatment of mass spectra in [27] recognises that their distribution assumes quark-hadron duality. However, the true mass spectra will have contributions with $X_{s/d} = V$, where V is a vector resonance. Of these only the lowest lying resonances, $K^*(892), \rho$ and ω are narrow enough to significantly distort the spectrum shape and potentially invalidate the assumption of quark hadron duality. For this reason the hadronic mass is split into two regions, a low mass resonant region and a high mass continuum of states. This is illustrated in the right hand plot of

[4] Although the parameters are essentially independent there is a naive expectation that μ_π^2 increases as m_b decreases. Previous *BABAR* analyses therefore make correlated systematic variations to reflect this and this analysis maintains that convention.

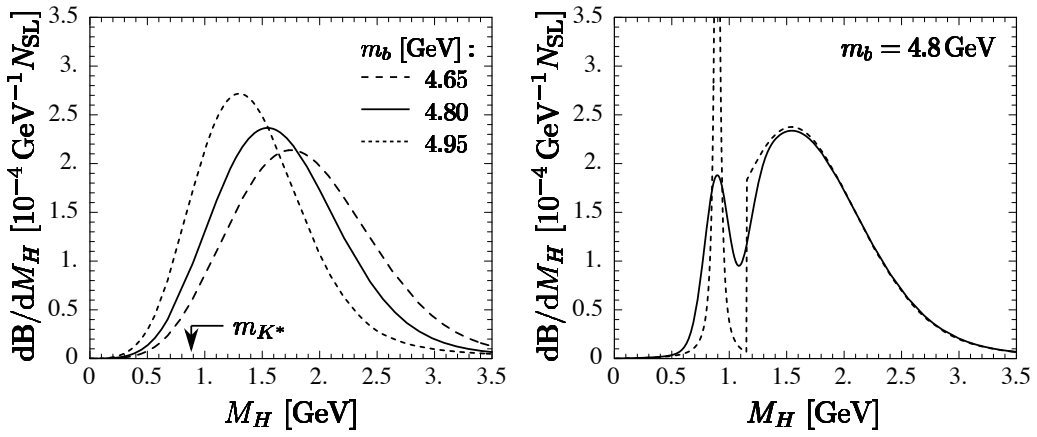


FIGURE 1.12: Plots taken from Figure 8 of [27] showing normalised hadronic mass spectra for $B \to X_s\gamma$ decays. The authors of this paper use the symbol M_H in place of m_X which is used in the text of this thesis. The left plot shows variations in m_b with $4.65\,\mathrm{GeV}/c^2$ the long-dashed line, $4.8\,\mathrm{GeV}/c^2$ the solid line and $4.95\,\mathrm{GeV}/c^2$ the short-dashed line. The right plot shows a combined mass spectrum separating the low mass resonant region from the high mass region where there is a continuum of states. The relevance of the dashed and solid lines are described in the text.

figure 1.12 for $B \to X_s\gamma$ decays. The K^*[5] peak is modelled as a relativistic Breit-Wigner (BW) distribution[6] and the continuum of states is introduced at a threshold, m_X^T, as illustrated by the dashed line. The solid line shows this spectrum model convoluted with a Gaussian distribution of width $\sigma = 100\,\mathrm{MeV}/c^2$, to emulate potential experimental resolution effects of hadronic mass reconstruction. The model used for the analysis presented here emulates the dashed line with m_X^T=1.0 GeV/c^2 and assumes that to a good approximation such a mass spectrum does not invalidate the assumption of quark-hadron duality [27].

1.3 Experimental Measurements of $|V_{td}/V_{ts}|$

The following section briefly summarises the existing experimental measurements of $|V_{td}/V_{ts}|$ from both B mixing and radiative penguin decays.

1.3.1 Measurement from B Mixing

Figure 1.13 shows the dominant Feynman diagram contributing to neutral B meson mixing. As with radiative penguin transitions the more massive t quark dominates the

[5]Unless otherwise stated K^* refers to the $K^*(892)$ resonance.

[6]The BW distribution has the form $|\mathcal{A}(s)|^2 = m_0^2\Gamma^2/[(s-m_0^2)^2 + (m_0\Gamma^2)]$ where s is the square of the centre of mass energy, m_0 is the resonance pole and Γ the resonance width.

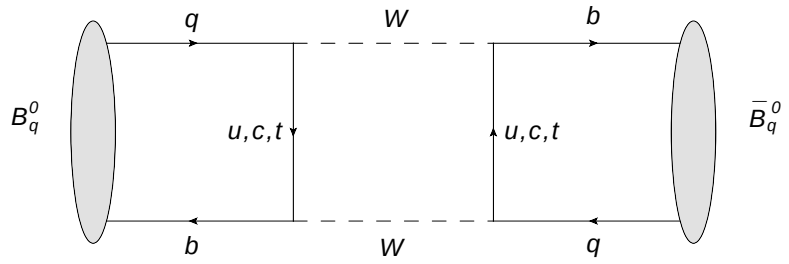


FIGURE 1.13: Dominant Feynman diagram contributing to neutral B meson mixing. The B meson spectator quark $q = d, s$.

virtual fermion exchange. In $B_d^0 \to \overline{B}_d^0$ oscillation (where $B_d^0 = |\bar{b}d\rangle$) the charged weak current has contributing factors from V_{td} and V_{td}^*. In the transition $B_s^0 \to \overline{B}_s^0$ (where $B_s^0 = |\bar{b}s\rangle$) these factors are substituted for V_{ts} and V_{ts}^* respectively. It can be shown that [14],

$$\left|\frac{V_{td}}{V_{ts}}\right|^2 = \xi_{QCD}^2 \frac{\Delta m_d}{\Delta m_s} \frac{M_{B_s^0}}{M_{B_d^0}}, \tag{1.76}$$

where ξ_{QCD}^2 is a theoretical factor calculated using Lattice QCD techniques [14] and Δm_q and $M_{B_q^0}$ are the respective mixing frequency and mass of the B_q^0 system. The CDF collaboration measure Δm_s=17.77±0.12 ps $^{-1}$ and combining this with the world average Δm_d=0.507±0.005 ps $^{-1}$ they calculate $|V_{td}/V_{ts}| = 0.209 \pm 0.001 \pm 0.006$ [32] where the first error is experimental and the second error theoretical.

1.3.2 Measurements from Radiative Penguin Decays

The *BABAR* and BELLE collaborations have measured the exclusive radiative penguin transitions $B \to \rho\gamma$ [33, 34]. As discussed above such measurements can be combined with world average $B \to K^*\gamma$ measurements [14] to extract $|V_{td}/V_{ts}|$. The *BABAR* measurements give $|V_{td}/V_{ts}|$=0.233±0.025±0.022 [33] and BELLE measure $|V_{td}/V_{ts}|$= 0.195±0.020±0.015 [34]. In both cases the first error is experimental and second error theoretical. The BELLE result has an improved theoretical error as their result takes advantage of a more recent theoretical calculation. The previous version of this analysis measures $|V_{td}/V_{ts}|$=0.177±0.043±0.001 [1].

1.3.3 Summary of $|V_{td}/V_{ts}|$ Measurements

To date the measurement from neutral B meson mixing is clearly more accurate, with almost negligible experimental error and relatively small theoretical error from Lattice QCD calculations. Although less competitive the radiative penguin measurements are

currently consistent with the B mixing result; exclusive decay modes have similar experimental and theoretical errors whereas the previous version of this analysis is limited by the experimental error. The aim of the analysis presented in this thesis is to significantly reduce the experimental error of the previous measurement.

Chapter 2

The *BABAR* Experiment

The *BABAR* experiment collected e^+e^- annihilation data from the interaction region of the energy-asymmetric PEP-II storage rings at the Stanford Linear Accelerator Center (now the SLAC National Accelerator Laboratory) in California, USA. Running from May 1999 until April 2008 it was designed to experimentally probe the theoretical description of heavy flavour physics in the SM. The primary physics goal from the analysis of *BABAR* data, as outlined in the Technical Design Report [35], is the study of *CP*-violating asymmetries in the decays of B^0 mesons to *CP* eigenstates. For example the measurement of CP asymmetries between the decay $B^0 \to J/\psi K^0_S$ and its charge conjugate constrains the SM parameter $\sin 2\beta$, where β is an angle of the Unitarity Triangle discussed in the previous chapter. Additional physics goals include measurements of B meson decays sensitive to different elements of the CKM matrix, with the ultimate aim of imposing redundant constraints on its parameters. The full dataset of just under $0.5\,\mathrm{ab}^{-1}$ also allows rare B decays with branching fractions of order 10^{-6} to be studied. Finally the clean environment and high luminosities of the *BABAR* experiment permits precision measurement of many τ and charm decays. Overall *BABAR* has a very rich physics program, more details of which can be found in the *BABAR* physics book [4] and from browsing any of the physics publications from the *BABAR* collaboration, which now number in excess of 350 papers.

This chapter provides a concise description of the PEP-II collider and *BABAR* detector. Emphasis is placed on B meson production and on the tracking, particle identification and calorimetry of both primary and secondary daughter particles from B meson decays as these aspects are the most relevant to the analysis described in this thesis. Technical specifications and performance requirements[1] of the relevant detector subsystems in each of these categories are discussed.

[1] Despite the *BABAR* experiment no longer being operational, in this thesis all design specifications and performance requirements are described in the present tense.

2.1 PEP-II Asymmetric e^+e^- Collider

The PEP-II asymmetric-energy two-storage-ring accelerator [36] collided bunches of electrons and positrons at a centre of mass (CM) energy of 10.58 GeV. This energy corresponds to the mass of the $\Upsilon(4S)$ [14], an S-wave $b\bar{b}$ bound resonant state which is the first of the excited $\Upsilon(nS)$ states above the $B\overline{B}$ threshold. The $\Upsilon(4S)$ accounts for almost 25% of the hadronic e^+e^- cross section at this energy and decays almost exclusively to $B\overline{B}$ pairs[2] [4]. The PEP-II design luminosity of $3 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ corresponds to a 3 Hz production rate of B meson pairs, making it a suitable laboratory for the study of their decays. Due to the low Q-value of the decay $\Upsilon(4S) \to B\overline{B}$, B mesons in the CM frame are produced almost at rest. In order to accurately measure the flight length difference, and hence decay time differnece of a $B\overline{B}$ pair, their decay vertices must be well separated in the laboratory frame. For this reason 9.0 GeV electrons are collided with 3.1 GeV positrons head-on resulting in the $\Upsilon(4S)$ obtaining a Lorentz boost of $\beta\gamma = 0.56$ in the laboratory frame.

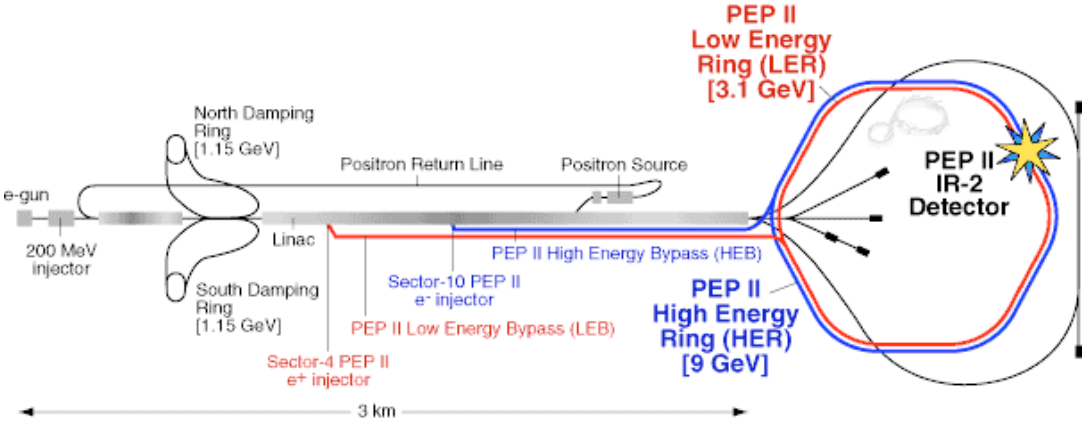


FIGURE 2.1: Schematic diagram of the PEP-II storage rings and linear accelerator.

Figure 2.1 shows a schematic representation of the PEP-II storage rings and the linear accelerator (linac) from which electrons and positrons were injected. Electrons bunches, produced from an electron gun at the end of the linac, are accelerated to around 1 GeV and fed into a damping ring where they are condensed to higher densities required for collision. Positrons are created from colliding further electrons accelerated to around 30 GeV with a tungsten-rhenium target and then collected in bunches. These are also accelerated to around 1 GeV and fed into a separate damping ring. From the damping rings both electrons and positrons are accelerated along the linac, using RF cavities

[2] $B\overline{B}$ refers to both $B^0\overline{B}^0$ and B^+B^- pairs.

powered by klystrons, to energies of 9.0 GeV and 3.1 GeV respectively. They are then injected into the PEP-II storage rings: the electron beam, travelling clockwise, into the so-called high energy ring (HER) and the positron beam, travelling anti-clockwise, into the corresponding low energy ring (LER). Both beams are brought into collision at the interaction region (IR) surrounded by the *BABAR* detector. The beams are focused with sets of quadrupole magnets before colliding in the IR and then steered away from each other using dipole magnets. Instantaneous luminosity measurements are made by detecting radiative Bhabha photons emitted by incoming positrons [37]. PEP-II achieved a record instantaneous luminosity of $12.1 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, which is more than four times the design.

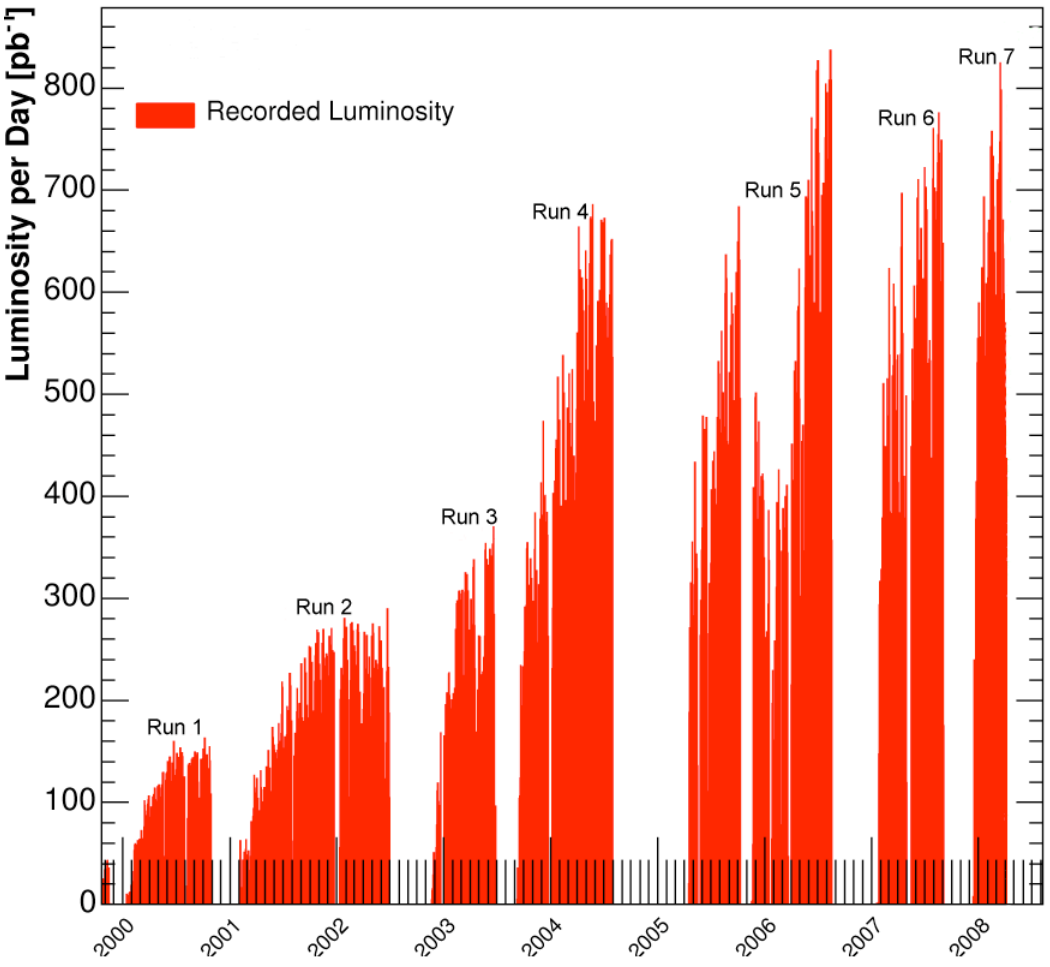


FIGURE 2.2: Plot showing the integrated luminosity per day of the PEP-II collider for the duration of its running period.

The total integrated luminosity of data collected by *BABAR* at the IR of the PEP-II collider is $531.4\,\mathrm{fb}^{-1}$, a detailed breakdown of which is shown in figure 2.2. Of these data $432.9\,\mathrm{fb}^{-1}$ were recorded at the $\Upsilon(4S)$ CM energy, with an additional $45.3\,\mathrm{fb}^{-1}$

collected at an energy just below the resonance to allow simulation independent studies of hadronic backgrounds from $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$) processes, which in this thesis are referred to as 'continuum' events. Data collection periods are divided into distinct intervals each known as a 'Run'. There are a total of seven data collection Runs which are indicated on figure 2.2. The time between running periods was used to perform both detector and accelerator maintenance and upgrades. Data from Runs 1-6 were collected at or just below the $\Upsilon(4S)$ resonance and so form the full dataset used for the analysis described in the subsequent chapters. Run 7 collected the remaining $53.2\,\text{fb}^{-1}$ of data at or just below both the $\Upsilon(2S)$ and $\Upsilon(3S)$ resonances and was ended with an 'energy scan' above the $\Upsilon(4S)$ resonance, it therefore does not form part of the analysis dataset used in this thesis.

2.2 The *BABAR* Detector

The *BABAR* detector is comprehensively described in [38], a summary of which is given here. It is typical of most modern detectors in high energy physics in that it consists of a series of sub-detectors arranged in a nested cylindrical onion-like structure around the IR. Each sub-detector is dedicated to measuring a particular characteristic of the long-lived daughter particles produced from initial interactions. Here long-lived refers to either stable particles (i.e. electrons, protons and photons) or those whose average lifetime is sufficiently long (e.g. muons, positrons[3], charged pions and charged kaons) that there is a high probability they interact with one or more of the *BABAR* sub-detectors. Generally each sub-detector design is optimised to measure the properties of a subset of long-lived particles. Ultimately these measured properties are used to infer the character of primary particle decays and hence extract the physics of interest from an event.

Figure 2.3 shows longitudinal and transverse cross-sections of the *BABAR* detector. Each of the sub-detectors are labelled and in order of increasing distance from the IR these consist of: a silicon vertex tracker (SVT) and drift chamber (DCH) for charged particle detection; a detector of internally reflected Cherenkov radiation (DIRC) used in charged particle identification; an electromagnetic calorimeter (EMC) for the detection of photons and identification of electrons; and an instrumented flux return (IFR) for muon detection. A superconducting solenoid, surrounding all sub-detectors with the exception of the IFR, produces a 1.5 T uniform magnetic field along the principal axis of the drift chamber for the measurement of charged track momenta. The collision axis is offset from this axis by about 20 mrad to minimise the effect of the solenoid field on the colliding beams. In figure 2.3 it can be seen that the *BABAR* detector itself has an

[3]Positrons are of course stable particles, however, in the presense of the matter of the *BABAR* detector there is a high probability they will annihilate.

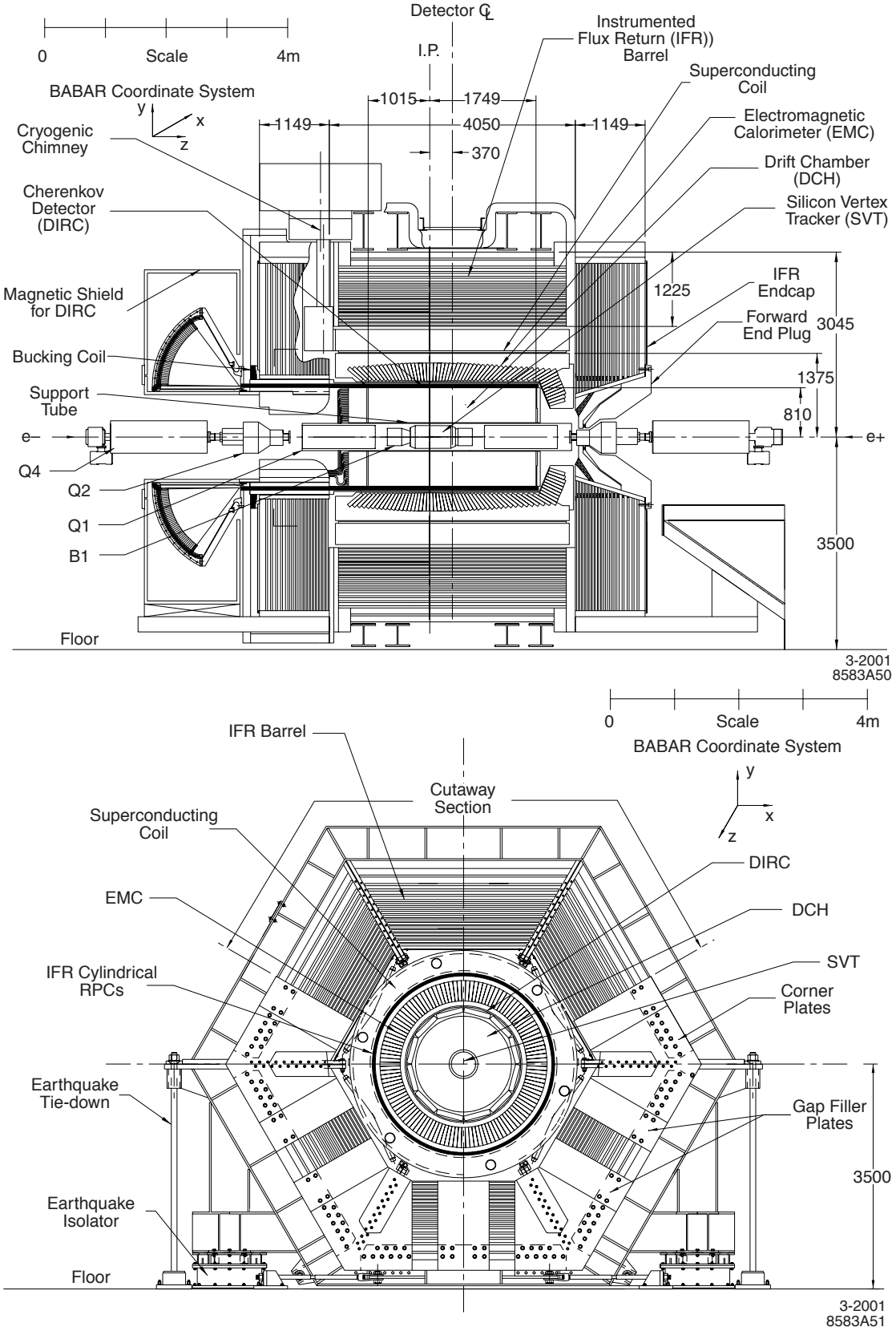


FIGURE 2.3: The *BABAR* detector shown in (top plot) longitudinal cross-section (bottom plot) transverse cross-section. In this figure the IR is denoted IP.

asymmetric layout, offset from the beam-beam interaction point. Due to the Lorentz boost of the $\Upsilon(4S)$ in the laboratory frame, this asymmetry maximises the geometric acceptance of the detector in the CM frame. The coordinate system of the *BABAR* detector is defined as a right handed Cartesian system; the z axis is coincident with the drift chamber principal axis, the y axis points upwards and the x axis points away from the centre of the PEP-II storage rings. In spherical coordinates *BABAR* uses the American convention of (r, θ, ϕ) as the distance, zenith and azimuth, respectively.

2.2.1 Design Goals

Physics studies using data from the *BABAR* experiment generally require full reconstruction of B meson decays with multiple charged and neutral secondary particles. Often there is a need to identify the flavour of unreconstructed B mesons produced in an event. Consequently the *BABAR* detector and hence each of the sub-detectors are required to fulfill a number of strict design goals [35]. The detector needs a uniform acceptance down to small zenith angles in the CM frame and a high reconstruction efficiency for long-lived charged particles with momenta as low as $60\,\mathrm{MeV}/c$ and for photons with energy as low as $20\,\mathrm{MeV}$. The resolution of a single B decay vertex is required to be around $100\,\mu\mathrm{m}$ both transverse and parallel to the z axis to ensure that the reconstructed decay vertices of both B mesons are sufficiently separated from each other in the laboratory frame. An efficient and accurate particle identification (PID) method for charged particles is required over a wide range of momenta. This must be sufficient to separate electrons and muons with a low probability of them being misidentified as hadrons. It must also allow for accurate identification of different hadron species, again with low misidentification probabilities.

2.2.2 Detection of Charged Particles

Generally when detecting the properties of long-lived charged particles one aims to measure the particle four-momentum, production vertex and trajectory followed while in the detector volume. This measurement should be non-destructive, meaning the energy transfer between the particle being measured and the detecting medium should be minimal. The *BABAR* detector has two sub-detectors for the measurement of such properties, a silicon strip vertex detector surrounded by a gas-filled wire drift chamber.

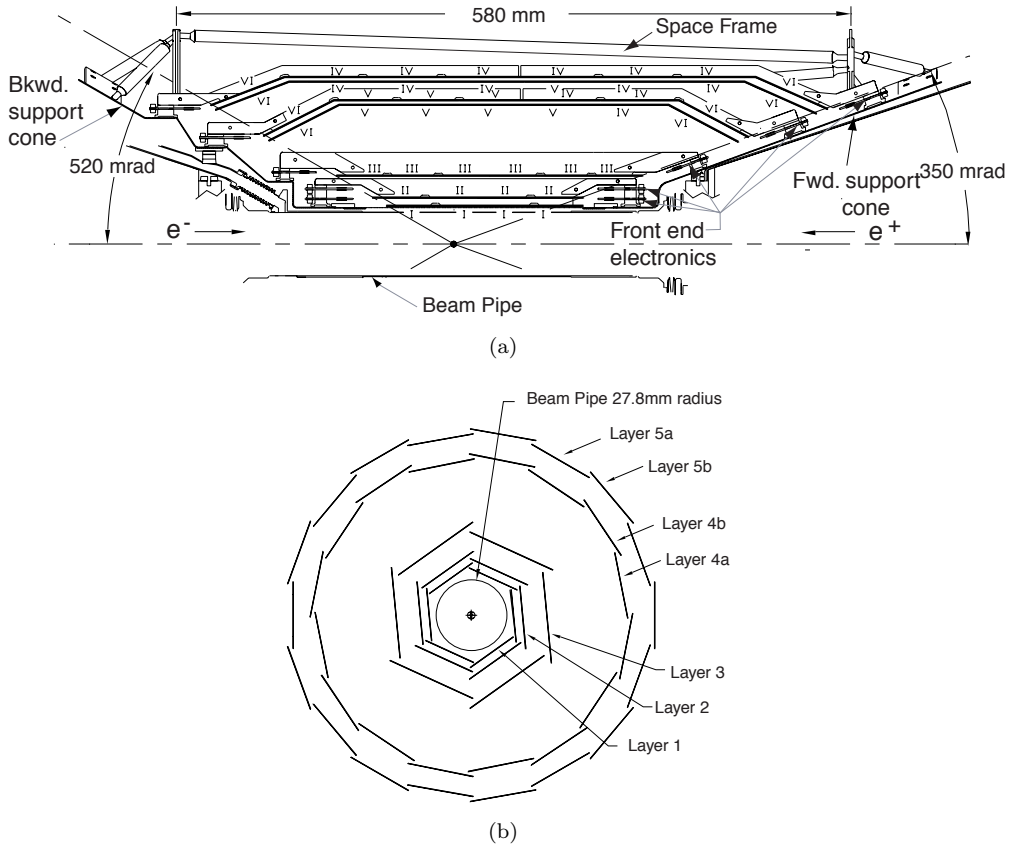


FIGURE 2.4: Schematic (a) longitudinal and (b) transverse sections of the *BABAR* silicon strip vertex detector.

2.2.2.1 The Silicon Vertex Tracker

The SVT is designed to give high spatial resolution of charged particle trajectories near the IR, thus providing the required resolution of primary particle decay vertices in a collision event. It consists of five layers of double-sided silicon strips and surrounds the 5.6 cm diameter beryllium beampipe which itself corresponds to 1.06% of a radiation length (X_0). Signals are characterised by charge deposits created from ionisation in the doped silicon bulk as a charged particle traverses the detector. Longitudinal and transverse sections of the SVT are shown in figure 2.4. For each layer the inner strip of a module is orientated orthogonally to the outer strip. The former are oriented transversely to the principal axis and measure the position in z of any charged particle traversing them; the latter are parallel to the principal axis and measure the position in ϕ. From figure 2.4 it can be seen that the inner three layers, which each contain six modules, are axial in z and tilted in the transverse plane, forming an overlap which ensures full radial coverage. These inner layers provide the impact parameter measurement of a charged

particle trajectory. The outer two layers, which in order of increasing radius contain 16 and 18 modules, are arched in z resulting in less material for solid angle coverage (i.e. less than if these layers had been entirely axial in z) and increasing the crossing angle of charged particles near the edge of acceptance. In the transverse plane the outer layers are divided into sub-layers at slightly different radii, again to ensure complete coverage in this plane. These layers are important in the reconstruction of trajectories of charged particles with low transverse momentum and for associating higher momentum trajectories with corresponding trajectories in the DCH. The SVT has a solid angle coverage of 90% in the CM frame and the material in the tracking volume is around 4% of X_0.

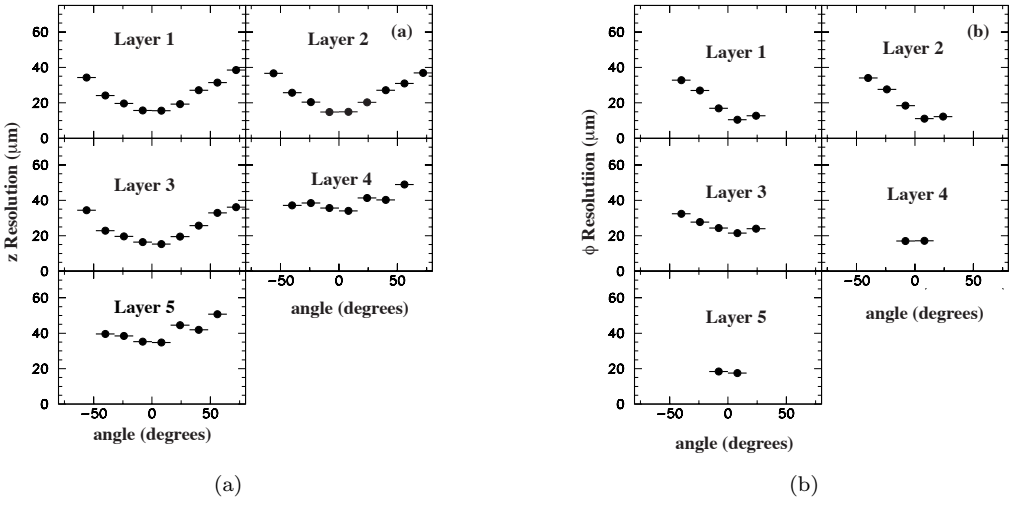


FIGURE 2.5: Measured resolutions in (a) z and (b) ϕ as a function of incident track angle for charged particles traversing the SVT. The performance for each of the five layers is shown.

Figure 2.5 shows the spatial resolution performance in z and ϕ for the SVT as a function of incident track angle. It can be seen that small resolutions ranging between around 15 μm and 50 μm are achieved in z depending on layer and incident track angle, with the corresponding resolution for ϕ ranging from around 10 μm to 35 μm. For the inner layers an asymmetry in the resolution of ϕ can be seen; this is due to the tilted modules in the transverse plane. The smaller incident range seen in the resolution of ϕ in the outer layers is also attributable to the geometry of these layers.

2.2.2.2 The Wire Drift Chamber

The DCH is designed for the efficient and high precision measurement of the momentum and angles of charged particles with momenta from around 120 MeV/c to 4 GeV/c and

higher. It provides the reconstruction of any decay vertices outside the SVT, such as $K_S^0 \rightarrow \pi^+\pi^-$ decays, and thus requires a longitudinal resolution of around 1 mm. The average momentum of a charged particle at *BABAR* is less than 1 GeV/c which means multiple scattering results in significant limitations on trajectory resolution. The DCH consists of 40 cylindrical layers of hexagonal drift cells and is filled with a 4:1 helium:isobutane gas mixture. The helium minimises multiple scattering and has a short drift time, while the isobutane absorbs any photons produced by ionised particles to prevent secondary ionisations from the photoelectric effect. Each cell contains a gold-plated tungsten-rhenium sense wire, held at a typical operating voltage of 1960 V and surrounded by six grounded gold-plated aluminium field wires. Charged particles traversing a cell ionise the gas resulting in a charge avalanche on the sense wire. At operating voltages the typical gain in a cell from an avalanche is around 5×10^4. The hexagonal layout ensures a close to circular field across the majority of a cell.

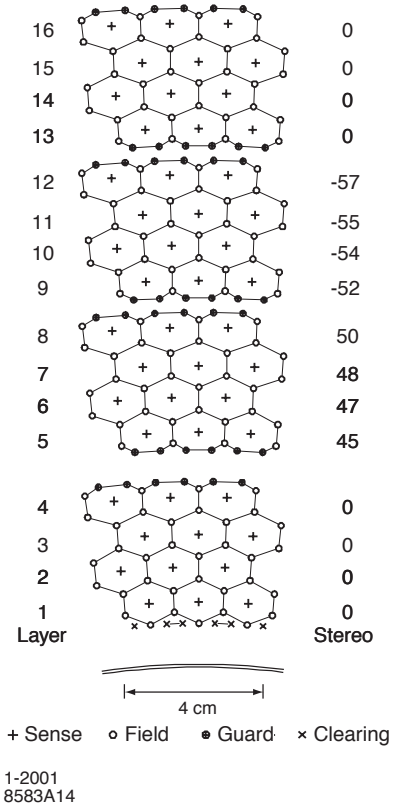


FIGURE 2.6: Schematic representation of the first four superlayers in the DCH. The stereo angle in mrad shows the AUVA structure used for longitudinal measurement.

The 40 layers of the DCH are split into 10 'superlayers', each of four layers. The first four superlayers are shown in figure 2.6. To obtain longitudinal position measurements six of the 10 superlayers are tilted relative to the z axis though a given stereo angle. The stereo angles alternate between axial (A) and stereo (U,V) pairs where U has a positive

stereo angle and V negative. The resultant order of superlayers is AUVAUVAUVA. Transversely the gas mixture and wires of the DCH correspond to 0.2% of X_0; this increases to 1.08% when the inner and outer support walls are included.

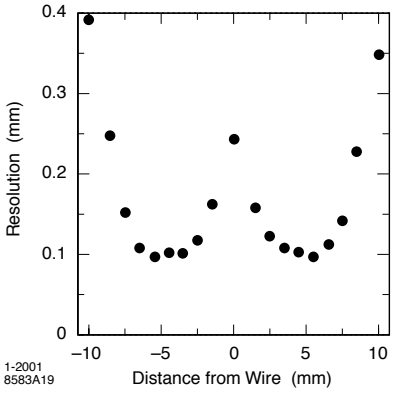


FIGURE 2.7: The average resolution of cells in layer 18 of the DCH as a function of distance from the sense wire.

The electronic readout from the DCH is designed to measure the drift time and integrated charge for every wire with a signal. Primary ionisation cluster positions are determined from timing of the leading edge of the amplified signal with a designed position resolution of 140 μm averaged over cells in the transverse plane. Figure 2.7 shows the average position resolution for multi-hadron events as a function of drift distance in layer 18 for tracks to the left and right of a sense wire. The resolution deteriorates closer to the wire due to insufficient charge build up from the avalanche and again further from the wire due to timing uncertainties resulting from the diffusion of the avalanched charge. The average resolution for the cell is close to design.

2.2.2.3 Charged Particle Trajectory Reconstruction

Both the SVT and the DCH are immersed in a uniform 1.5 T magnetic field directed along the z axis from the superconducting solenoid. Due to the resultant Lorentz force a charged particle will follow a helical path in the transverse plane, the radius of curvature (RoC) of which gives a measurement of the transverse momentum magnitude, p_T. The direction of the RoC determines whether the particle holds a positive or negative charge. A high resolution measurement of the momentum of a charged particle therefore requires an accurate reconstruction of its trajectory. *BABAR* achieves this through complimentary measurements from the SVT and DCH. Each detector makes a position measurement in a single active element of its volume known as a 'hit'. Reconstructed trajectories, known as 'tracks', are created from pattern recognition software which performs a Kalman fit

on a collection of individual hits to determine if they are consistent with the passage of a single charged particle. The fit algorithm corrects for measured inhomogeneities in the magnetic field, although systematic uncertainties remain as the field was mapped in the absence of PEP-II magnets. Reconstructed tracks are described by five parameters, $(d_0, z_0, \phi_0, \tan\lambda, \omega)$, all measured at the point of closest approach (POCA) to the z axis; d_0 and z_0 are the distances from the POCA to the transverse plane and the z axis respectively, ϕ_0 is the azimuthal angle, λ the dip angle relative to the transverse plane, and ω is the track curvature given as $1/p_T$.

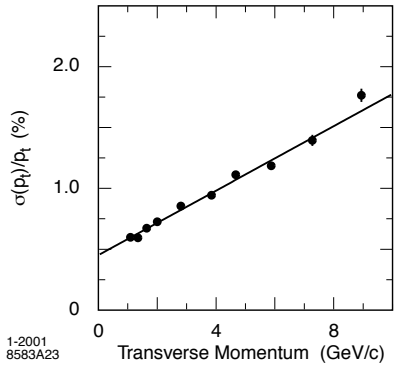


FIGURE 2.8: The combined resolution of the SVT and DCH in transverse momentum for cosmic muons.

Figure 2.8 shows the combined p_T resolution of the SVT and DCH for cosmic muons. The resolution is clearly linear in p_T and the corresponding fitted line gives a resolution function,

$$\sigma_{p_T}/p_T = (0.13 \pm 0.01)\% \times p_T + (0.45 \pm 0.03)\%. \tag{2.1}$$

Tracks which pass the Kalman fitter selection are available through the *BABAR* computing framework for event reconstruction in a physics analysis.

2.2.3 Charged Hadron Identification

The SVT and DCH measure spatial momentum of reconstructed charged particle tracks. However, to fully reconstruct the four-momentum the mass of that particle must also be known. The *BABAR* detector is required to differentiate between the properties of different charged hadrons, in particular pions, kaons and protons. PID for long-lived charged leptons is discussed in section 2.2.5.

Ionisation energy deposited in the tracking sub-detectors is particularly important for low momentum tracks. Characteristic Bethe-Bloch curves [14] show distributions of ionisation energy deposited by a charged particle, $\mathrm{d}E/\mathrm{d}x$, as a function of momentum

and can thus indicate which particle traversed the tracking medium. Figure 2.9 shows

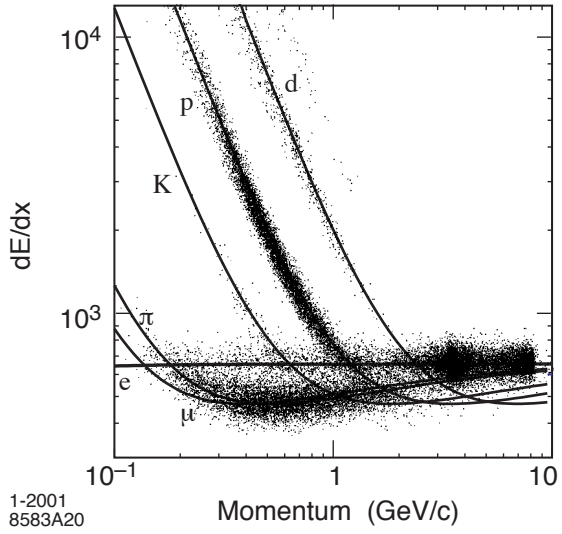


FIGURE 2.9: $\mathrm{d}E/\mathrm{d}x$ as a function of track momentum comparing Bethe-Bloch predictions to data for various particle types. The unit of the $\mathrm{d}E/\mathrm{d}x$ curve is arbitrary due to corrections for the deposited charge in individual DCH cells.

the measurement of $\mathrm{d}E/\mathrm{d}x$ in the DCH as a function of momentum and compares data to the Bethe-Bloch hypothesis for different particle types. It can be seen that for low momentum tracks in the $1/\beta^2$ region a good separation exists between kaons, pions and protons. As particles reach the minimum ionisation energy and approach the relativistic limit the separation worsens. The resolution of the DCH in $\mathrm{d}E/\mathrm{d}x$ is around 7% which gives good separation of pions and kaons up to momenta of 700 MeV/c. PID measurements from the tracking system are complementary to those from the Cherenkov detector (described below); however, charged particles with p_T <180 MeV/c will not reach the DIRC so the tracking volume provides the only measure of PID. Furthermore tracks with even lower transverse momentum will only deposit energy in the SVT which has a $\mathrm{d}E/\mathrm{d}x$ resolution of around 14%. This allows a 2σ separation of kaons and pions up to 500 MeV/c in this sub-detector alone.

Generally the PID algorithms for charged tracks at *BABAR* combine relevant information from different sub-systems into custom PID selectors which maximise the probability of correct identification. These selectors are available to analysts during event reconstruction and are described in section 3.2.5.

2.2.3.1 The Cherenkov Detector

It can be seen from figure 2.9 that the ability of the DCH to distinguish between charged particle species with momenta above 700 MeV/c deteriorates rapidly. Many physics processes studied with *BABAR*, including the analysis presented in subsequent chapters, require good separation between reconstructed pions and kaons up to momenta of 4 GeV/c. It is for this reason that *BABAR* employs a sub-detector dedicated to charged particle PID, the DIRC. Charged particles with velocity $\beta = p/E$ will emit Cherenkov radiation in a cone of opening angle $\cos\theta_c = 1/\beta n$ if $\beta > 1/n$, where n is the refractive index of the medium being traversed. It follows that given the particle momentum and angle, which in *BABAR* is given by the tracking system, a measurement of θ_c will give the particle mass. The DIRC is intended to exploit this principle but must avoid degrading the energy resolution of the EMC which lies beyond it. For this reason the detection mechanism for Cherenkov photons is outside the active detector volume.

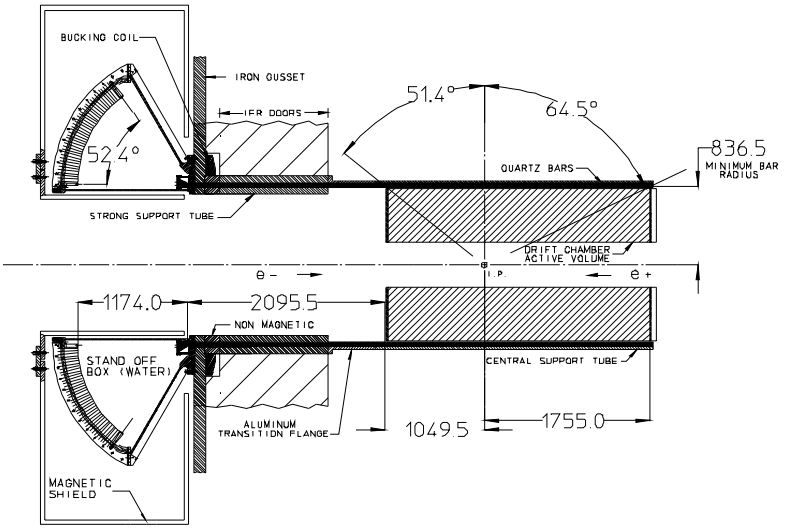


FIGURE 2.10: Longitudinal section of the DIRC showing the active medium and external detection system.

The DIRC is shown in longitudinal section in figure 2.10. It consists of 144 fused silica bars, with n=1.473, which act as the radiative material and also transfer the emitted photons by total internal reflection to the rear end of the detector, preserving θ_c (forward going photons are reflected by a mirror). Here they enter a stand-off box (SOB) filled with purified water whose refractive index is close to that of the quartz bars to reduce internal reflection and refraction at the interface. The SOB is mounted with 10,752 photomultiplier tubes to detect the Cherenkov light which forms a conic section whose opening angle is θ_c, once corrected for the effects of refraction.

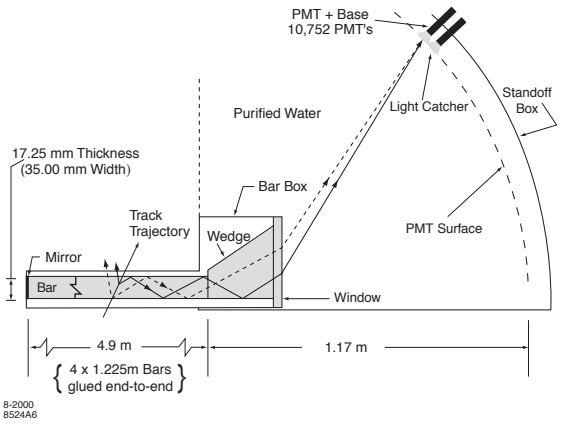


FIGURE 2.11: Schematic showing the principal of operation for the DIRC.

Figure 2.11 shows a schematic demonstrating the operating principle of the DIRC as described above. It should be noted that this schematic is an oversimplification showing the operation in only two dimensions of what is a three dimensional system. Transversely

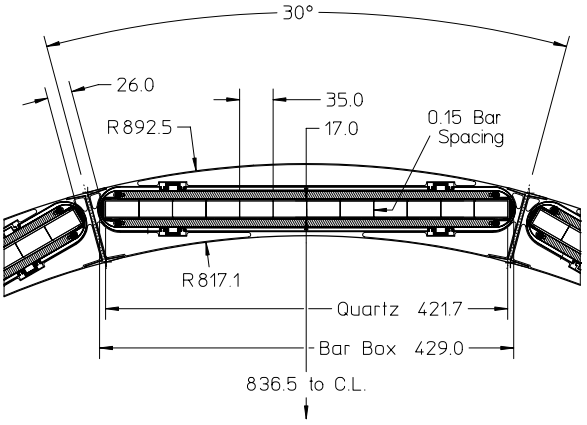


FIGURE 2.12: A transverse section of one of the 12 polygonal sides of the DIRC barrel.

the DIRC bars are arranged in a 12-sided polygonal barrel, with each side containing 12 bars as shown in figure 2.12. The material in the detector volume corresponds to 17% of X_0 and due to the polygonal structure the acceptance in the azimuth is around 94%. In the CM frame the acceptance is around 83% of the polar angle cosine.

Figure 2.13 (a) shows the expected separation of pions and kaons as a function of momentum for $B^0 \to \pi^+\pi^-$ events inferred from the measured Cherenkov angle resolution and number of Cherenkov photons per track in dimuon events. It can be seen that the separation varies from greater than 10σ at $2\,\mathrm{GeV}/c$ to less than 3σ at $4\,\mathrm{GeV}/c$. Figure 2.13 (b) shows the efficiency and misidentification probability using the DIRC for selection of charged kaons from a $D^0 \to K^+\pi^-$ control sample. The average kaon selection

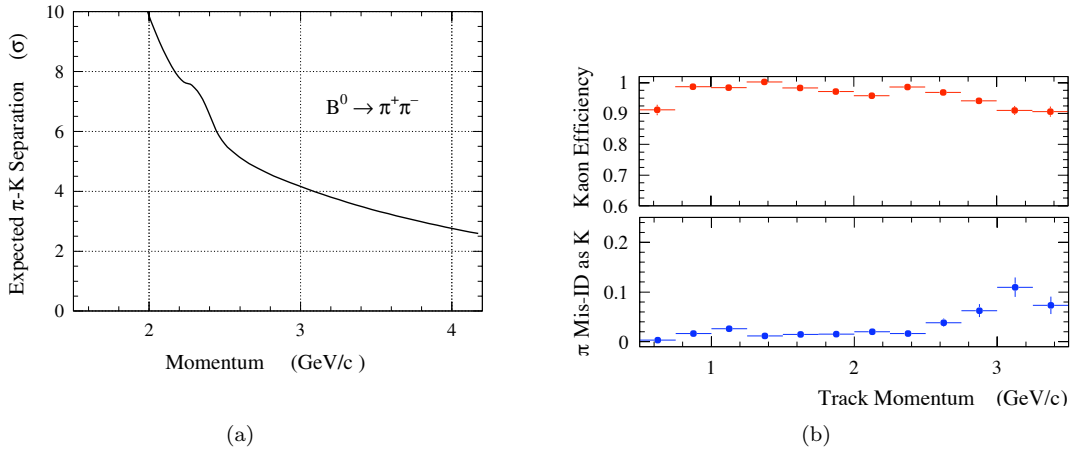


FIGURE 2.13: DIRC expected separation of pions and kaons for (a) $B^0 \to \pi^+\pi^-$ events and (b) the measured kaon efficiency and pion misidentification in $D^0 \to K^+\pi^-$ events.

efficiency and pion misidentification are 96.2±0.2% and 2.1±0.1% respectively, where the errors are statistical.

2.2.4 Photon Detection

Of particular importance to many physics studies with *BABAR* data is the accurate measurement of both the angle and energy of photons. Such measurements are required to be highly efficient with excellent resolution of photon energies ranging from 20 MeV to 9 GeV. The lower bound is set by the requirement for efficient reconstruction of both $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ decays. The upper bound relates to the need to measure QED processes such as $e^+e^- \to \gamma\gamma$ and $e^+e^- \to e^+e^-(\gamma)$ for calibration and luminosity determination. In the presented analysis photons with energies of up to 4.5 GeV in the laboratory frame are required for event reconstruction and many of the B meson decays to be considered require the reconstruction of at least one $\pi^0 \to \gamma\gamma$ or $\eta \to \gamma\gamma$ decay. The measurement of photon energies and angles at *BABAR* is carried out with an electromagnetic calorimeter.

2.2.4.1 The Electromagnetic Calorimeter

The EMC is a hermetic total absorption calorimeter finely segmented with 6580 thallium-doped caesium iodide (CsI(Tl)) scintillating crystals. Geometrically it is split into a cylindrical barrel and forward endcap giving it full coverage in the azimuth and 90% solid angle coverage in the CM frame. The barrel contains 5760 crystals arranged into 48 rings of 120 crystals, while the endcap has eight rings in total of which, in order of decreasing zenith angle, three contain 120 crystals, three contain 100 crystals and two

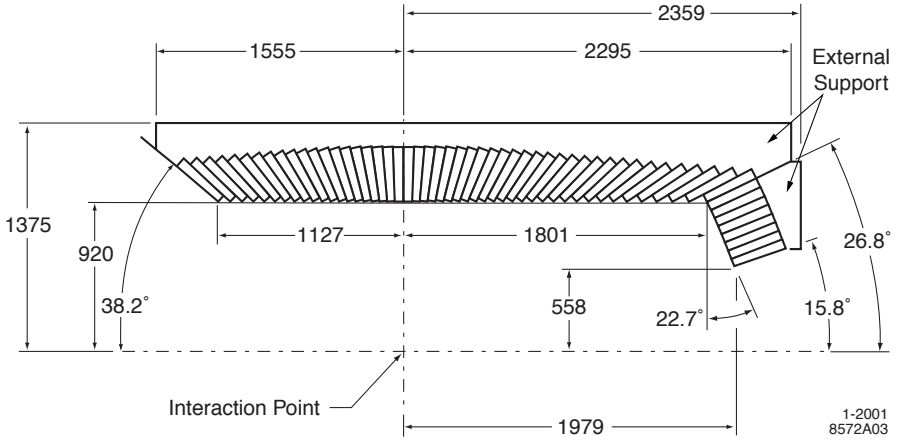


FIGURE 2.14: A longitudinal section showing the top half of the EMC which indicates the crystal layout in the barrel and endcap sections. All dimensions are in mm.

contain 80 crystals. CsI(Tl) has a high light yield and small Molière radius giving high energy and angular resolution as well as a X_0 of 1.85 cm to ensure shower containment. Crystal lengths vary from 16 X_0 in the backward barrel region at high zenith to 17.5 X_0 in the forward barrel and endcap to minimise leakage from higher energy particles moving in the boost direction.

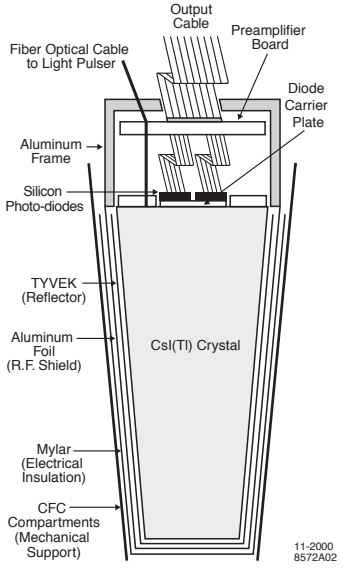


FIGURE 2.15: Schematic representation of a CsI(Tl) scintillating crystal used in the *BABAR* EMC.

As shown in figure 2.15 crystals are trapezoidal in shape. They guide the scintillating light, which is proportional to the deposited energy, through total internal reflection from highly polished surfaces to a pair of silicon diodes used for detection. To reduce transmission an additional two layers of diffuse white reflector surround the crystal edges.

The front face area of crystals varies to achieve hermicity but is typically $4.7\times4.7\,\mathrm{cm}^2$, close to the Molière radius to maximise the angular resolution while limiting the required number of crystals. Readout diodes are mounted on the back face of each crystal via an optical epoxy to maximise light transmission. Signals are amplified and digitised in the detector volume and then passed out via fibre optical cables.

2.2.4.2 Calorimeter Resolution

The energy resolution of a homogeneous crystal calorimeter is given by summing the following terms in quadrature,

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt[4]{E/\mathrm{GeV}}} \oplus b, \tag{2.2}$$

where E and σ_E are the respective energy and corresponding RMS error of a detected photon. The a term is an energy dependent variable primarily describing fluctuations in photon statistics, but is also affected by noise from both beam-generated background and electronics. The b term, which dominates at energies $>1\,\mathrm{GeV}$, accounts for non-uniformity in light collection, leakage and absorption in the detector and calibration uncertainties.

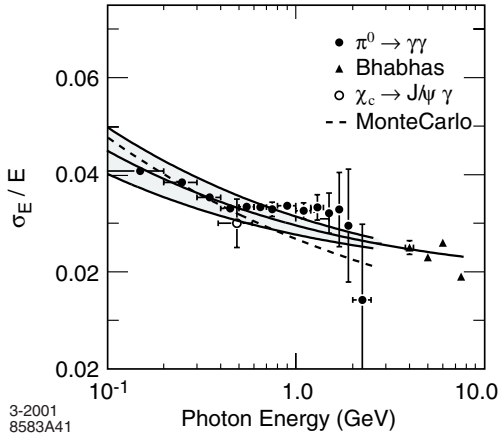


FIGURE 2.16: EMC energy resolution in data for different processes. The central solid line represents the fit used to extract the resolution measurement while the shaded region represents $\pm1\sigma$ bands.

The EMC measures low energy resolutions directly with a radioactive source[4] giving $\sigma_E/E = 5.0 \pm 0.8\%$ at $6.13\,\mathrm{MeV}$. High energy resolutions are calculated from Bhabha scattering events, where energy can be predicted from the e^- zenith angle, and gives $\sigma_E/E = 1.9 \pm 0.1\%$ at $7.5\,\mathrm{GeV}$. Figure 2.16 shows EMC energy resolutions for a variety

[4]Through the reaction $^{16}O^* \to^{16} O + \gamma$.

of processes. The fitted resolution gives $a = 2.32 \pm 0.30\%$ and $b = 1.85 \pm 0.12\%$ which is in reasonable agreement with Monte Carlo studies of the expected resolution.

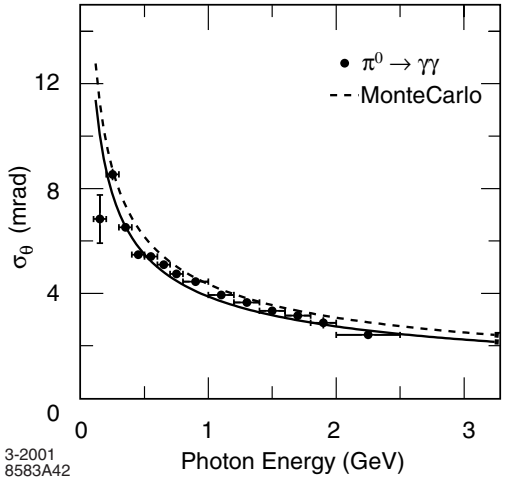


FIGURE 2.17: EMC angular resolution as measured from $\pi^0 \to \gamma\gamma$ decays. The solid line shows the fit used to extract the resolution.

The angular resolution depends on the transverse crystal size and distance from the IR and is empirically parameterised as the sum of an energy dependent term, c, and a constant term, d:

$$\sigma_\theta = \sigma_\phi = \frac{c}{\sqrt{E/\mathrm{GeV}}} + d. \tag{2.3}$$

Angular resolution is measured from π^0 and η decays to photons of approximately equal energy. Figure 2.17 shows the results for π^0 decays with angular resolution as a function of energy. The fitted line is of the empirical form of (2.3) and gives $c = 3.87 \pm 0.07\,\mathrm{mrad}$ and $d = 0.00 \pm 0.04\,\mathrm{mrad}$ which is slightly better than expected from Monte Carlo studies.

2.2.4.3 Calorimeter Energy Clusters

Electromagnetic showers in the EMC are typically spread over many crystals. Pattern recognition software is used to efficiently identify these clusters and then perform a search for individual local maxima known as 'bumps'. This differentiates between clusters from a single source and those formed from overlapping showers. Clusters are required to have at least one crystal with an energy deposit greater than 10 MeV and adjacent crystals are included if they have an energy above 1 MeV or if they have a neighbour with energy exceeding 3 MeV. The total energy of a cluster must be in excess of 20 MeV. A weighted iterative algorithm calculates the energy associated with each bump in a cluster and the respective angular position is calculated with a centre-of-gravity method

using logarithmic weights. Finally reconstructed tracks are projected onto the inner face of the EMC to determine if a particular bump centroid is associated with a charged particle. This process is known as 'track-matching'; if no associated tracks are found the bump is assumed to originate from a neutral particle. For all clusters a correction is made to the measured energy to account for energy leakage due to the gaps between crystals.

Many useful parameters can be calculated for a bump of n crystals in the EMC [39]. Two of these quantities are directly relevant to the analysis in this thesis.

The lateral moment, L_{bump}, is defined as,

$$L_{bump} = \frac{\sum_{i=3}^{n} E_i r_i^2}{\left(\sum_{i=3}^{n} E_i r_i^2\right) + (E_1 + E_2) r_0^2}, \tag{2.4}$$

where E_i is the energy of the ith crystal[5] in the bump, r_i is the distance between the corresponding crystal and bump centroid and r_0=5 cm is the average distance between two crystal front-faces. Electromagnetic showers (e.g. from an electron or photon) tend to deposit a large fraction of their energy within a few crystals whereas hadronic showers typically have a larger energy spread. It follows that hadronic showers will on average have a larger lateral moment.

The second moment, S_{bump}, is defined as,

$$S_{bump} = \frac{\sum_i E_i \Delta\alpha_i^2}{\sum_i E_i}, \tag{2.5}$$

where $\Delta\alpha_i$ is the angle between the crystal and the bump centroid. Merged π^0 decays with a single local maximum in the EMC will generally have an elliptical energy distribution whereas a high energy photon will have an energy distribution symmetric about the centroid. The energy deposit from a high energy photon will therefore have a lower second moment associated with it.

2.2.5 Electron and Muon Identification

The identification of long-lived charged leptons is important for determining the flavour of semileptonic B^0 and $\overline{B}^0$ meson decays, for the reconstruction of τ decays and for the study of QED processes such as $e^+e^- \to \mu^+\mu^-$. The tracking system and DIRC can provide some information for such PID as discussed above. Electron PID can also make use of EMC information; generally an electron will deposit all of its energy in an EMC

[5]Crystal numbering is energy ordered such that the most energetic crystal is labelled 1, the next energetic crystal labelled 2 and so on.

shower and due to its relatively low mass the ratio E/p will be close to one, which is not the case for more massive charged hadrons. Muons tend to pass through the EMC as a minimum ionising particle (MIP) and traverse the steel solenoid flux return. To assist in the separation of charged hadrons from muons the flux return is instrumented to detect any ionisation from traversing muons.

2.2.5.1 The Instrumented Flux Return

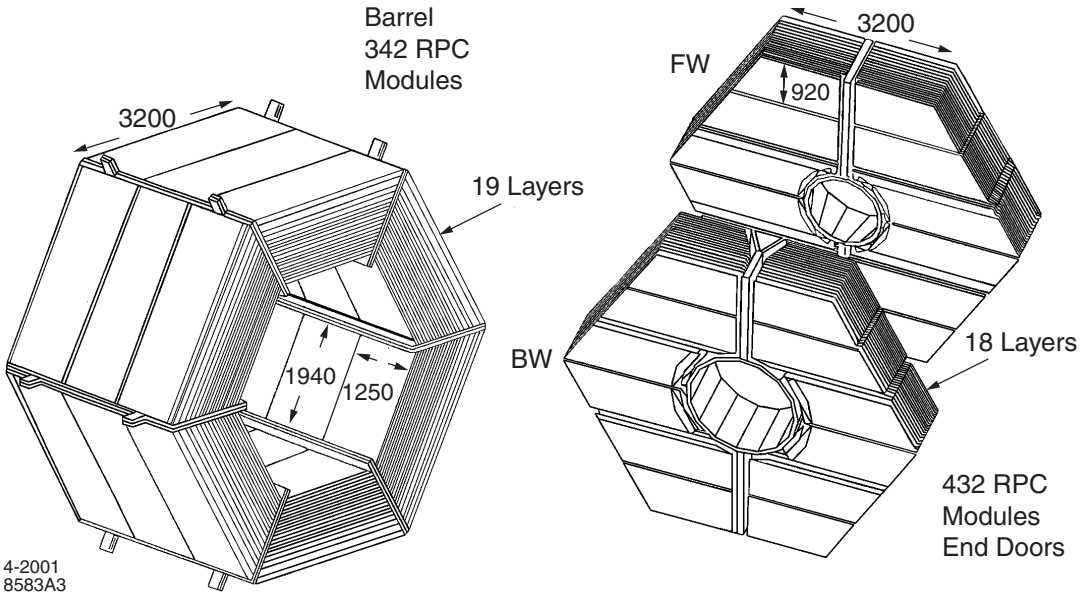


FIGURE 2.18: Layout of the IFR showing the segmented steel and instrumented layers for the barrel and two end doors.

The IFR makes use of the *BABAR* steel flux return as a muon filter and hadron absorber. Gaps between the finely segmented sections of the steel are instrumented to detect streamers from ionising particles. Figure 2.18 shows the layout of the IFR which consists of a hexagonal barrel with 19 layers of instrumentation and two end doors with 18 layers. The depth of steel layers varies from 2 cm at the inner layers to around 10 cm at the outer layers. Initially the IFR was instrumented with 806 single gap resistive plate chambers (RPCs) with 57 in each of the six barrel sectors and 108 in each of the four half end doors. An additional two cylindrical layers of 32 RPCs were installed between the EMC and solenoid to detect particles exiting the EMC.

Figure 2.19 shows a schematic cross-section of an IFR RPC. The active volume is filled with a mixture of argon, freon and isobutane which creates a discharge on the passage of an ionising particle. The discharge is detected via two capacitive readout strips which are placed orthogonally to measure both ϕ and z of the ionisation. In the first years

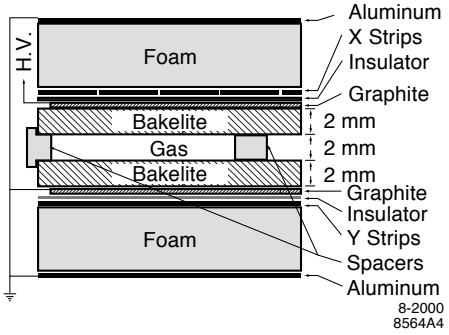


FIGURE 2.19: Section showing the principal RPC design.

of operation a serious degradation in RPC detection efficiency was observed due to a number of factors including construction flaws. This led to the barrel instrumentation being replaced with limited streamer tubes (LSTs) [40]. Of the six barrel segments, RPCs in the top and bottom were replaced between Run 4 and Run 5 while those in the remaining four segments were replaced between Run 5 and Run 6. The end doors were instrumented with RPCs for the lifetime of the experiment although some modules were replaced with a slightly modified design.

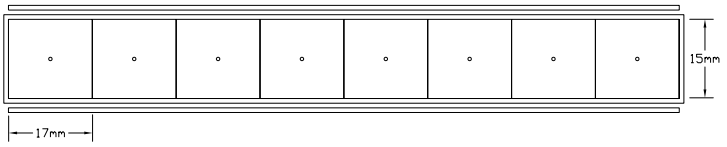


FIGURE 2.20: Section showing the principal LST design.

Figure 2.20 shows an IFR LST sketched in transverse cross section. Each cell has a single gold-plated anode wire at high voltage, typically around 5500 V, and contains a gas mixture of CO_2, argon and isobutane (to prevent secondary ionisations) in a (89:3:8) ratio. Ionising particles cause a discharge in the gas which is read out from the wire, giving a ϕ coordinate of the discharge. A simultaneous charge is induced on a plane mounted below the tube consisting of 96 copper strips perpendicular to the wire direction which read out the z coordinate along the 4 m length of the plane. The first LST installation fitted 24 z-planes and 388 tubes, while the second stage installed 48 z-planes and 776 tubes.

Muon detection efficiency as a function of the pion rejection efficiency in the IFR barrel for high energy muons is shown in figure 2.21. The deterioration in RPC performance between 2000 and 2004 can clearly be seen. LST segments installed for Run 5 significantly improve the muon detection efficiency compared to the remaining barrel RPCs.

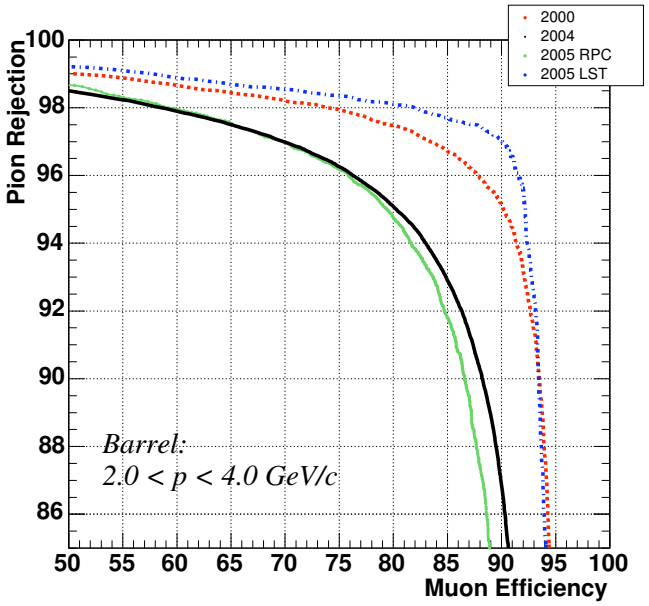


FIGURE 2.21: Muon detection efficiency as a function of pion rejection efficiency for high energy muons in the IFR barrel. The dotted line shows RPC efficiencies in 2000, the upper and lower solid lines show the RPC efficiencies in 2004 and 2005 respectively. The dotted-dashed line shows the efficiency for LSTs installed for Run 5 in 2005.

In their first year of operation the LSTs even out-performed the original efficiency of the RPCs.

2.2.6 The Trigger and Data Acquisition System

The *BABAR* trigger is designed to manage the readout rate of detector sub-systems, whose latency buffers store signals from all e^+e^- interactions. The required rate is determined by the bandwidth at which data can be transferred and stored before a more detailed offline reconstruction and analysis can be performed. The trigger employs a two level hierarchy: a 'Level 1' hardware trigger (L1T) with input from the DCH, EMC and IFR and a 'Level 3' software trigger (L3T) which performs real time event reconstruction and classification. The combination of this hierarchy is designed to be over 99% efficient in the identification and readout of $B\overline{B}$ events as well as at least 95% efficient for continuum events and 90-95% efficient for $\tau^+\tau^-$ events. The system is also designed to be redundant in its event identification to allow the accurate study of trigger efficiencies.

2.2.6.1 The Hardware Trigger

The L1T monitors a constant stream of information from the DCH and EMC in the drift chamber trigger (DCT) and electromagnetic calorimeter trigger (EMT) respectively. The IFR also provides information to allow for triggering of cosmic ray and $\mu^+\mu^-$ events, mainly for diagnostics. The L1T is entirely digital using reduced data derived from the respective subsystems.

Input to the DCT consists of one bit for each of the 7104 DCH cells conveying time information derived from the sense wire of that cell. Groups of hits in adjacent DCH cells are identified by a track segment finder. This information is passed to a binary link tracker which groups segments into tracks. Tracks recognised by the DCT are then used to form trigger primitives depending on whether they satisfy predefined criteria including reaching a given superlayer, exceeding a p_T threshold or satisfying a cut in z.

The EMT divides the EMC into 280 towers in a 7×40 array in (θ, ϕ) of which the barrel contributes 6×40 towers and the endcap 1×40. Each tower in the barrel consists of 24 crystals in a 8×3 array, while the endcap towers are formed of wedges in ϕ containing between 19 and 22 crystals. The energy from each tower in θ is summed independently for the 40 ϕ sectors. Neighbouring ϕ energy strips are then added together to identify any showers which may cross two adjacent ϕ-sectors. The resultant energy for all 40 strips are compared to predefined thresholds to form trigger primitives. The sums are then ORed between neighbouring ϕ sectors giving a 20-bit ϕ map for each primitive.

Primitives from the DCT and EMT are passed in parallel to the global level trigger (GLT) as bitmaps in ϕ. The GLT compares these bitmaps to 24 predefined physics signatures. If the GLT finds a match it passes this information to the Fast Control and Timing System (FCTS) which performs the trigger decision. Depending on the GLT output the FCTS can issue a 'Level 1 Accept' (L1A) or 'prescale' the trigger. A L1A results in the latency buffers for all sub-systems being read out for further processing. Prescaled triggers will only issue a L1A for for every nth event which pass their selection criteria. The priority of the L1T is to issue L1As for multihadron events which correspond to physics events of interest such as B meson decays. Events such as Bhabha scattering used for luminosity calculations can be prescaled as this introduces no bias. The GLT contains both pure DCT and pure EMT L1As allowing unbiased efficiency measurements to be made for both of these trigger subsystems. The DCT and EMT are independently up to 99% efficient in identifying $B\overline{B}$ events, demonstrating the redundancy in the L1T. Output rates for L1As are around 2.5 kHz under typical

running conditions with the L1T contributing less than 1% of deadtime[6] in the *BABAR* data acquisition system.

2.2.6.2 The Software Trigger

Events passing a L1A are processed in real time with an online computing farm and passed through the L3T. All event information is available to the L3T allowing a more sophisticated event selection process than the L1T, such as better track reconstruction and EMC energy clustering. Event filters select and classify events of interest using predefined scripts which contain selection criteria for different physics processes. The L3T reduces the output rate by around an order of magnitude compared to the L1T. Events passing the L3T filters are written to a temporary event store before being passed to offline computing farms which perform a more complete event reconstruction allowing for detailed quality monitoring before data are made available for analysis.

[6]Deadtime refers to periods when new data cannot be read out due to previous data being processed.

Chapter 3

Event Selection

Collision events recorded by the *BABAR* detector are passed through a number of general and custom software filters in order to extract physics signatures of interest. This chapter describes the selection techniques used to maximise $B \to X_{s/d}\gamma$ [41] event extraction while efficiently rejecting combinatoric backgrounds. An overview of the analysis method is given and the variables used to distinguish between signal and background events are defined. Finally the development, optimisation and implementation of custom event filters are described.

3.1 Analysis Overview

Making a fully inclusive measurement of $B \to X_{s/d}\gamma$ decays can provide a significant challenge due to their rarity. The relative high energy of the photon provides a powerful

Mode	$B \to X_d\gamma$	$B \to X_s\gamma$
1	$B^0 \to \pi^+\pi^-\gamma$	$B^0 \to K^+\pi^-\gamma$
2	$B^+ \to \pi^+\pi^0\gamma$	$B^+ \to K^+\pi^0\gamma$
3	$B^+ \to \pi^+\pi^-\pi^+\gamma$	$B^+ \to K^+\pi^-\pi^+\gamma$
4	$B^0 \to \pi^+\pi^-\pi^0\gamma$	$B^0 \to K^+\pi^-\pi^0\gamma$
5	$B^0 \to \pi^+\pi^-\pi^+\pi^-\gamma$	$B^0 \to K^+\pi^-\pi^+\pi^-\gamma$
6	$B^+ \to \pi^+\pi^-\pi^+\pi^0\gamma$	$B^+ \to K^+\pi^-\pi^+\pi^0\gamma$
7	$B^+ \to \pi^+\eta\gamma$	$B^+ \to K^+\eta\gamma$
8	$B^+ \to \pi^+\pi^0\pi^0\gamma$	$B^+ \to K^+\pi^0\pi^0\gamma$
9	$B^0 \to \pi^+\pi^-\pi^0\pi^0\gamma$	$B^0 \to K^+\pi^-\pi^0\pi^0\gamma$

TABLE 3.1: Reconstructed exclusive final states. The $B \to X_s\gamma$ final states are obtained by substituting a charged pion for a charged kaon in the $B \to X_d\gamma$ final states. All $\pi^0 \to \gamma\gamma$ and all $\eta \to \gamma\gamma$.

experimental signature which simplifies the task of identifying signal decays. Indeed for $B \to X_s\gamma$ it is sufficient to detect the high energy photon without reconstructing the X_s hadronic state [42, 43]. However, due to the nearly identical photon spectrum, such techniques are impractical when measuring $B \to X_d\gamma$ decays due to the dominance of the X_s sample. Instead this analysis reconstructs a subset of nine exclusive modes whose contributions are to be added, therefore making a semi-inclusive measurement. Consequently model-dependent corrections for unreconstructed contributions must be made to extract the fully inclusive branching fraction (BF) and hence $|V_{td}/V_{ts}|$. Table 3.1 lists the exclusive modes reconstructed[1] for $B \to X_d\gamma$ and $B \to X_s\gamma$ decays. The X_s hadronic states are defined by substituting a charged pion for a charged kaon in the corresponding X_d hadronic state with the kaon charge determined by the flavour of the parent B meson. Identical cuts are used for both samples leading to many of the experimental systematic errors cancelling in the ratio of branching fractions. Due to the excess of expected events in the CKM favoured X_s sample all cuts and filters are optimised to provide the most statistically significant measure of the X_d sample. Final state π^0 and η mesons are reconstructed exclusively from the decays $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ which corresponds to 98.8%(39.4%) of the $\pi^0(\eta)$ decay fraction [14].

The analysis is divided into two bins of reconstructed hadronic mass: a low mass bin, $0.5 \leq m_X < 1.0\,\mathrm{GeV}/c^2$, and a high mass bin, $m_X \geq 1.0\,\mathrm{GeV}/c^2$. This division is motivated by the distinct characteristics expected for signal decays in each mass bin, as discussed in section 1.2.3.2. Consequently simulated signal events used for analysis must reflect this distinction. The low mass $X_d(X_s)$ region is dominated by the resonant transition $B \to \rho/\omega\gamma(B \to K^*\gamma)$ and is therefore modelled exclusively with these decays. Resonant particle masses are modelled as Breit-Wigner distributions using the world average width for that particle [14]. The high mass bin is modelled as a cocktail of non-resonant decays according to the KN photon spectrum [27] using m_X^T=1.0 GeV/c^2. The hadronic state is generated through phase space decays from JETSET [44]. This model does not account for the vector resonance contributions with masses greater than $1.0\,\mathrm{GeV}/c^2$ such as high end tails of the K^* and ρ distributions or the measured $B \to K_2^*(1430)\gamma$ transition [14]. It is assumed that the low mass resonance tails are to a good approximation accounted for in the JETSET phase space distribution; the variation of hadronic final states due to higher mass resonances is considered as a source of systematic uncertainty in section 5.2.2. The upper limit of the high mass bin is chosen such that the non-resonant measurement be as inclusive as possible while considering the increase in combinatoric backgrounds at higher values of m_X. This is primarily constrained by computational limitations implementing the pre-reconstruction event filters described in

[1]The choice of reconstructed modes is based on those used in the previous version of this analysis: modes 1-7 [1]. Modes 8 and 9 were added here with the aim of making a more inclusive measurement; however, their addition was found to be a disadvantage in the statistical optimisation discussed below.

section 3.4. Post-reconstruction the choice of upper mass limit forms part of the cut optimisation procedure using simulated data, discussed in section 3.6.7.

Overall there are four experimental measurements: the semi-inclusive BF obtained from the sum of exclusive reconstructed states in the low and high mass bins for $B \to X_d\gamma$ and $B \to X_s\gamma$ transitions. The signal yields of these transitions are extracted from experimental data using a two dimensional extended maximum likelihood fit which is described in chapter 4. The measured BFs are then used to estimate total BFs in each mass region using a model dependent extrapolation. The extrapolated BFs are combined for each flavour to give fully inclusive BFs, the ratio of which are then used to extract a value for $|V_{td}/V_{ts}|$ according to the theoretical relations discussed in chapter 1.

3.1.1 Simulated and Experimental Data Samples

This analysis uses $429.1(44.8)\,\text{fb}^{-1}$ of experimental data recorded with CM energy at ($\sim$40 MeV below) the $\Upsilon(4S)$ resonance. Data recorded at the $\Upsilon(4S)$ corresponds to $(470.9 \pm 2.8) \times 10^6$ $B\overline{B}$ pairs [45]. The analysis is optimised using Monte Carlo (MC) simulated data and then applied to the experimental data, thus eliminating bias which can arise from basing the event selection and distribution modelling on real data. The MC data used in this analysis at least match the size of the experimental dataset. In the *BABAR* simulation framework particle decays are generated from a database of allowed decay modes, each with a given BF, and the daughter particle kinematics are determined from a phase space model of the decay. Responses of the detector subsystems are modelled using the GEANT4 package [46] and QED radiative corrections are modelled with PHOTOS [47].

To optimise custom event filters for $B \to X_d\gamma$ signal extraction, simulations of both signal events and events which can contribute to combinatoric backgrounds are required. Table 3.2 lists the different simulated datasets used for analysis optimisation, showing the total number of generated events. The first nine MC datasets in table 3.2 are signal decays[2] whereby one generated B meson is forced to decay to the final state indicated and the other can decay to any allowed final state. Signal decays do not directly correspond to the reconstructed final states listed in table 3.1; they contain all allowed final states for the modelled transition, i.e. all allowed decays of the resonant particle in the low mass region or all allowed phase space decays from JETSET in the high mass region. The reconstructed decays therefore form a subset of the signal MC data. Due to the

[2]Note that MC data modelling the high mass bin are generated separately for charged and neutral B meson decays. To distinguish these the hadronic component of the decay is labelled with a double subscript, the first component indicating the penguin transition quark flavour and the second component indicating the flavour of the B meson spectator quark.

Event Class	Simulated Transition	Simulated Events
KN and JETSET Signal	$B^+ \to X_{du}\gamma$	1,288,000
	$B^0 \to X_{dd}\gamma$	1,288,000
	$B^+ \to X_{su}\gamma$	11,178,000
	$B^0 \to X_{sd}\gamma$	11,178,000
Vector Resonance Signal	$B^+ \to \rho^+\gamma$	650,000
	$B^0 \to \rho^0\gamma$	650,000
	$B^0 \to \omega\gamma$	650,000
	$B^+ \to K^{*+}\gamma$	6,449,000
	$B^0 \to K^{*0}\gamma$	6,449,000
Generic Background	B^+B^-	708,762,000
	$B^0\overline{B}^0$	717,995,000
	$c\overline{c}$	1,128,544,000
	$q\overline{q}$ $(q = u, d, s)$	1,662,404,000

TABLE 3.2: Simulated datasets used for event selection optimisation showing the number of generated events before any cuts are applied.

Simulated Transition	Events $\geq m_X^T$	Fraction
$B^+ \to X_{du}\gamma$	1,201,936	0.933
$B^0 \to X_{dd}\gamma$	1,201,708	0.933
$B^+ \to X_{su}\gamma$	10,425,400	0.933
$B^0 \to X_{sd}\gamma$	10,429,063	0.933

TABLE 3.3: High mass bin simulated signal sizes after model based hadronic mass cut. The fraction of events remaining after the cut, relative to the full generated sample is also shown.

Simulated Transition	Events $< m_X^T$	Fraction
$B^+ \to \rho^+\gamma$	568,518	0.875
$B^0 \to \rho^0\gamma$	569,089	0.876
$B^0 \to \omega\gamma$	650,000	1.000
$B^+ \to K^{*+}\gamma$	5,752,171	0.892
$B^0 \to K^{*0}\gamma$	5,727,473	0.888

TABLE 3.4: Low mass bin simulated signal sizes after model based hadronic mass cut. The fraction of events remaining after the cut, relative to the full generated sample is also shown.

independent signal modelling in each mass bin, the boundary between signal samples is identified by placing the threshold cut of $m_X^T = 1.0\,\mathrm{GeV}/c^2$ on the generated hadronic mass. The number of simulated events for each sample after this cut are given in tables 3.3 and 3.4 for the high and low mass bins respectively. Normalisation of signal MC data to the *BABAR* recorded luminosity assume total inclusive BFs of $1 \times 10^{-5}(3.6 \times 10^{-4})$ for $B \to X_d\gamma(B \to X_s\gamma)$ decays. These BFs are taken from the expected order of magnitude of theoretical estimates in the case of $B \to X_d\gamma$ [22] and the world experimental average in the case of $B \to X_s\gamma$ [14]. The relative contributions of each vector resonance in the respective low mass bins are taken from world average BFs [14]. These resonant BFs are corrected to account for the fraction of events expected below m_X^T using the MC data fractions shown in table 3.4. For $B \to \rho/\omega\gamma(B \to K^*\gamma)$ transitions the calculated contribution corresponds to 10.2%(10.5%) of the above assumed inclusive widths.

Generic B pair events simulate other $\Upsilon(4S) \to B\overline{B}$ processes with both B mesons decaying to SM final states which are determined from a database of $B\overline{B}$ decays. The database essentially lists the BFs of known B meson decays and uses JETSET to generate higher multiplicity modes which have generally not been measured. These data include $B \to X_s\gamma$ events (but not $B \to X_d\gamma$ events) which must be vetoed to avoid double counting. Any signal or generic event with the decay of a neutral B meson pair models the effect of neutral meson mixing during event generation. Generic continuum events of light quark pairs from $e^+e^- \to q\overline{q}$ $(q = u, d, s, c)$ transitions use JETSET [44] to simulate the hadronisation of the quark pair. Events from $e^+e^- \to l^+l^-$ $(l = e, \mu, \tau)$ processes are found to contribute at a negligible level to combinatoric backgrounds and hence are ignored in MC data studies.

The simulated events produced for *BABAR* analyses are known not to model the data perfectly. To improve the agreement for neutral particles reconstructed with data from the EMC, asymmetric energy smearing and edge effect corrections are made [48]. This reduces systematic errors arising from differences between experimental data and MC events for such particles and is particularly important for $B \to X_{s/d}\gamma$ events where the photon deposits a large fraction of the total event energy in the EMC.

3.2 Event Reconstruction Framework

Data recorded by the *BABAR* detector which pass quality control checks are made available for analysis. Both experimental and MC data are regularly reprocessed to take advantage of improved central software filters. Each reprocessing cycle has a unique accessible dataset. Reconstruction software is interfaced through a custom designed central C++ analysis framework consisting of a number of software packages each dedicated to

to a subset of the reconstruction and which are called by the analyst as required. Collectively the packages are maintained and updated within numbered software releases. The analysis presented here uses reprocessed datasets labelled `R24a` with *BABAR* release number `24.3.6`, the most recent stable release for event reconstruction at the time of analysis. Detector conditions for a given running period and lookup tables for efficiency corrections in MC data are accessed from dedicated conditions databases. This analysis uses the conditions database `cond24boot09` which is the recommended database for the datasets and release number used.

Lists of fundamental reconstructed particles such as charged tracks and neutral clusters in the EMC are contained in a central event store for both real data and those simulated with GEANT4. All composite particles are reconstructed from these fundamental lists. Common composite particles, such as π^0 and η mesons, are reconstructed within the common analysis framework and are available at runtime. The following provides a summary of the lists used for reconstruction in this analysis. All cuts are made in the laboratory rest frame unless otherwise stated.

3.2.1 Charged Tracks

Charged tracks are taken from the list `GoodTracksLoose` which requires a momentum magnitude $\leq 10\,\text{GeV}/c$ and transverse momentum $\geq 0.05\,\text{GeV}/c$. Additionally these tracks must have a distance of closest approach to the IR of $|z| < 2.5\,\text{cm}$ along the principle axis and $< 1.5\,\text{cm}$ in the transverse plane. All tracks have an assumed pion mass hypothesis.

3.2.2 Photons

Photons are taken from the list `GoodPhotonLoose`. These consist of single bump EMC clusters without an associated charged track which have an energy deposit of $\geq 0.03\,\text{GeV}$ and lateral moment ≤ 0.8. A photon mass hypothesis is applied.

3.2.3 π^0 and η Mesons

Composite neutral meson candidates from dual photon decays are obtained from the lists `pi0DefaultMass` and `etaggDefault` for π^0 and η mesons respectively. These lists are constructed by adding the four-momentum of photon pairs from the `GoodPhotonLoose` list, with the additional requirement that photon energy deposits satisfy $0.03 \leq E_\gamma \leq 10.0 (0.05 \leq E_\gamma \leq 10.0)\,\text{GeV}$ and the invariant mass of photon pairs are in the range

$0.115 \leq m_{\gamma\gamma} \leq 0.150 (0.470 \leq m_{\gamma\gamma} \leq 0.620)\,\mathrm{GeV}/c^2$ for $\pi^0(\eta)$ meson candidates. Additionally η meson candidates are required to have a momentum magnitude in the range $0.2 \leq |p_\eta| \leq 10.0\,\mathrm{GeV}/c$ and π^0 candidates must have energy $\geq 0.2\,\mathrm{GeV}$. After satisfying these criteria the neutral candidates are refitted constraining the particle origin to the primary vertex (defined below) and their masses to the world average value [14]. This gives an unbiased improvement on the kinematic resolution of the particle which is limited due to detector resolution prior to refitting. Higher kinematic resolution is desirable when using the reconstructed neutrals to form composite candidates higher up the decay chain.

3.2.4 Vertex Fitting

Where composite candidates are reconstructed from two or more charged tracks fitting algorithms can determine the spatial coordinates of the origin of those tracks: the decay vertex of the composite particle. This analysis uses the *BABAR* fitting routine `Cascade` [49] whose goal is iterative χ^2 minimisation on the vertex hypothesis of the input particles. The `Cascade` fitter implicitly requires that momentum is conserved at a decay vertex, although particle momenta are allowed to vary within their measured errors in the fit. An additional geometric constraint requiring that all included tracks originate from the same spatial point can also be applied. Neutral particles, which are assumed to originate from the primary vertex, will have their momentum redefined so that they originate from the common decay vertex if they are included in the fit.

3.2.5 Charged Particle Identification Classifiers

As discussed in chapter 2 a number of variables, such as dE/dx in the tracking subsystems and Cherenkov angle in the DIRC, can be employed for the purpose of charged track PID. The usefulness of individual variables varies substantially for a given track flavour and momentum. To simplify the task of charged track PID *BABAR* employs a dedicated PID group who combine such variables into multivariate classifiers trained on well defined control samples. Generally the classifiers are accessed through PID selectors which give a binary output depending on whether the classifier output has passed a predefined cut. A single classifier can have multiple selectors depending on the tightness of cut on the classifier output, typically these are labelled `VeryLoose`, `Loose`, `Tight`, `VeryTight` and `SuperTight`. For each selector the PID group carry out momentum dependent efficiency and misidentification rate studies on real and simulated data. This analysis is primarily concerned with the discrimination between charged pions and kaons in order to separate $B \to X_s\gamma$ and $B \to X_d\gamma$ decays. For pion identification a multi-class

learning classifier, `pionKM`, is used [50]. This classifier is trained with 36 input variables from the SVT, DCH, DIRC and EMC using control samples of $D^{*+} \to D^0(K^-\pi^+)\pi^+$, $K_S \to \pi^+\pi^-$ and $\tau \to \nu_\tau 3\pi$ decays. For kaon identification a bagged decision tree classifier, `kaonBDT`, is used [50]. The classifier is trained on $D^{*+} \to D^0(K^-\pi^+)\pi^+$ control samples again with 36 input variables. The kaon classifier has an additional selector representing a looser cut than `VeryLoose`; this is labelled `NotAPion`. Control sample studies by the PID group show that the typical rate of charged kaons being misidentified as pions is less than 10% for all momenta.

3.2.6 Primary Vertex

The primary interaction in a e^+e^- collision event occurs at the PEP-II IR. The spatial position of this interaction can be determined from measurement of the beamspot. The size of the beamspot is typically 150 microns in x, 6 microns in y and 1 cm in z. For event reconstruction purposes a more accurate determination, particularly of the z coordinate, is required. This is achieved in the *BABAR* framework by fitting for the primary vertex on an event by event basis [49]. The fit combines all tracks with impact parameter in the transverse plane of less than 1 mm from the beamspot centre. For $B\overline{B}$ events the primary vertex will be shifted in the positive z direction from the e^+e^- interaction point due to the average B meson lifetime [14]. Essentially the primary vertex represents the mid-point between the two B meson decays. This is illustrated schematically in figure 3.1. In general neutral particle candidates within the *BABAR* framework have their four-momentum defined such that they originate from the primary vertex by default. The primary vertex resolution in a $B\overline{B}$ event is typically 100 microns in x and 115 microns in z. The y resolution from the primary vertex fit is around a factor of ten worse than the existing beamspot measurement, hence the beamspot measurement is used to constrain this coordinate.

3.3 Discriminating Variables

There are many variables which aid in the discrimination of reconstructed candidates from signal decay events compared to those originating from combinatoric backgrounds. These can broadly be separated into kinematic, topological and so-called 'tagging' variables. The variables considered in this analysis are defined here.

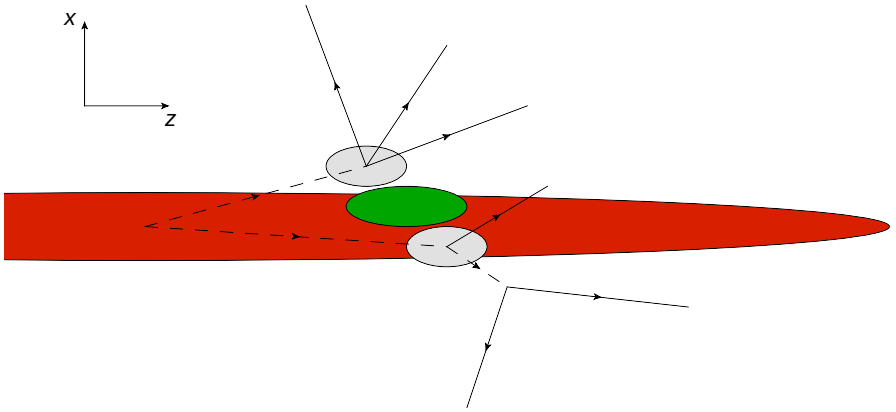


FIGURE 3.1: Schematic representation of the primary vertex for a $B\overline{B}$ event in the $x-z$ plane. The red ellipse represents the beamspot resolution. The green ellipse indicates the improved primary vertex resolution, particularly in the z direction. The grey ellipses show the respective B meson decay vertex resolutions. Note that the primary vertex is distinct from the $\Upsilon(4S)$ decay vertex. The resolutions shown are not to scale.

3.3.1 Kinematic Variables

In addition to placing kinematic cuts on the energy and momentum of reconstructed particles there are two kinematic variables of primary importance when considering B meson decays at *BABAR*. These are the beam energy substituted mass, m_{ES}, and the difference between the reconstructed and expected energy of the B meson, ΔE [51]. Explicitly they are defined as,

$$m_{ES} = \sqrt{\frac{1}{4}s - |\mathbf{p}_B^*|^2}, \tag{3.1}$$

$$\Delta E = E_B^* - \frac{1}{2}\sqrt{s}, \tag{3.2}$$

where $(E_B^*, \mathbf{p}_B^*)$ is the four-momentum of the reconstructed B meson and s is the Mandelstam variable for the transition $e^+e^- \rightarrow \Upsilon(4S)$, i.e. the CM energy squared. The asterisk superscript denotes a quantity in the CM frame. Both variables make use of the fact that the $\Upsilon(4S)$ resonance lies just above the $B\overline{B}$ threshold which results in each B meson having a CM energy half that of $\sqrt{s}$. It follows that ΔE will peak at zero for a correctly reconstructed B meson with an approximate Gaussian distribution resulting from detector resolution limits. The negative tail will be more pronounced due to energy leakage in the EMC which for a radiative decay such as the signal considered can be significant. The beam energy substituted mass is preferential to the reconstructed B mass as the beam energy is measured to a high degree of accuracy and does not suffer from detector resolution limitations inherent in the measurement of E_B^*. Furthermore an alternative mass hypothesis for charged tracks will not change the value of m_{ES}

whereas a correct measurement of E_B^* depends on the flavour of all tracks being correctly identified. For a correctly reconstructed B decay, m_{ES} will peak at the B mass, $m_B = 5.279\,\mathrm{GeV}/c^2$, with a Gaussian shape due to detector momentum resolution and the resolution in beam energy measurement. On average m_{ES} is observed to have an enhanced negative tail due to EMC energy leakage in the detection of neutral particles. To minimise the effect of energy leakage from the signal photon it is useful to define m'_{ES}, whereby the magnitude of $\mathbf{p}_B^*$ is recalculated constraining the photon energy such that ΔE=0, i.e.

$$E_\gamma^* + E_X^* - \frac{1}{2}\sqrt{s} = 0. \tag{3.3}$$

Here E_γ is the photon energy and E_X is the energy of the hadronic component of the decay. The potential introduction of a correlation between ΔE and m'_{ES} is discussed in section 4.1.3. Figures 3.2 and 3.3 show respectively simulated distributions for m'_{ES} (including an overlay of the corresponding signal m_{ES} distribution[3]) and ΔE for correctly reconstructed candidates and continuum background before any post-reconstruction cuts. The signal resolution is $\sim 50\,\mathrm{MeV}$ for ΔE and $\sim 5\,\mathrm{MeV}/c^2$ for m'_{ES}. The negative tail due to energy leakage can be seen in the signal distributions of both variables.

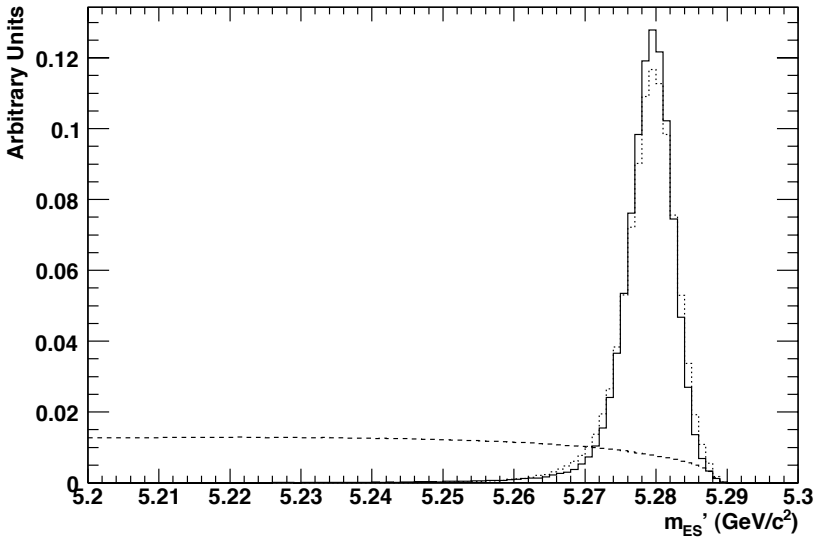


FIGURE 3.2: Normalised distributions of m'_{ES} for continuum (dashed line) and correctly reconstructed candidates (solid line) before post-reconstruction cuts. The dotted line overlays the distribution of m_{ES} for correctly reconstructed candidates to illustrate the resolution improvement obtained using m'_{ES}.

[3]Note that the improvement in resolution of m'_{ES} over m_{ES} is relatively minor. This is in contrast to early radiative penguin analyses at *BABAR* due to an improved energy leakage correction during EMC cluster reconstruction. However, to be consistent with earlier analyses m'_{ES} is generally preferred over m_{ES} as a discriminating variable.

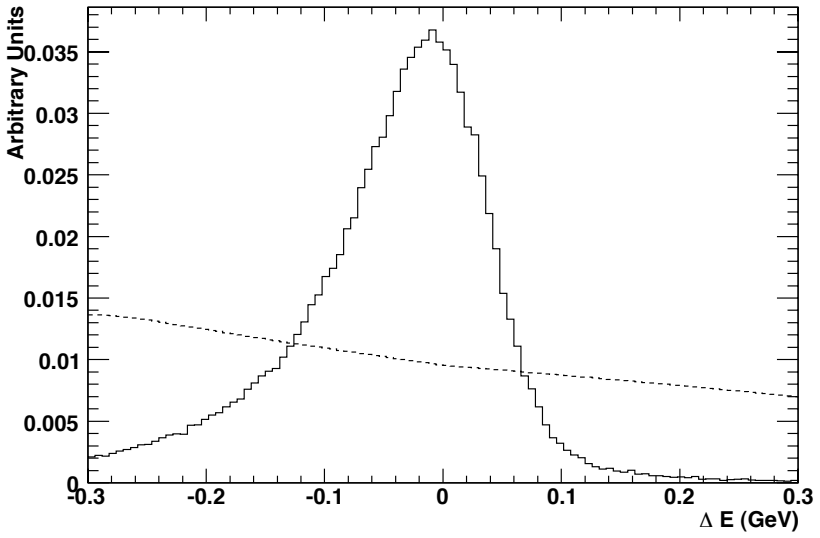


FIGURE 3.3: Normalised distributions of ΔE for continuum (dashed line) and correctly reconstructed candidates (solid line) before post-reconstruction cuts.

3.3.2 Topological Variables

Measuring the topology of detected particles in a given event can provide powerful discrimination between events arising from $\Upsilon(4S)$ decays and those arising from continuum backgrounds. The B mesons are produced almost at rest in the $\Upsilon(4S)$ frame and hence decay isotropically in that frame. Conversely particles produced in continuum events typically have large kinetic energies in the CM frame and form collimated jets along the axes of the initial quark and anti-quark pair. Therefore variables which distinguish the isotropic or jet-like nature of events can aid in the reduction of continuum backgrounds.

Continuum events containing reconstructed $B \to X_{s/d}\gamma$ candidates will have an associated high energy photon. Where this photon is a daughter of the decay chain from the initial quark hadronisation its direction will be highly correlated with other particles in the corresponding jet. Continuum background also arises from interactions involving initial state radiation (ISR), $e^+e^- \to q\bar{q}\gamma$. In this case the photon momentum will be largely independent of the individual jet directions. However, together both jets will be seen to recoil against the photon due to momentum conservation. It follows that the measurement of event shape variables with respect to the high energy photon can be used to further reduce continuum backgrounds.

In order to calculate many of the variables which can separate these event topologies it is necessary to distinguish between charged tracks and neutral clusters which were used to reconstruct the signal candidate and those which form the rest of the event

(ROE). For correctly reconstructed candidates the ROE is assumed to contain all particles detected from the other B meson decay in the event. All topological variables are calculated in the CM frame unless otherwise stated. In the following figures MC data distributions for correctly reconstructed candidates and continuum backgrounds before post-reconstruction cuts are presented.

3.3.2.1 Thrust

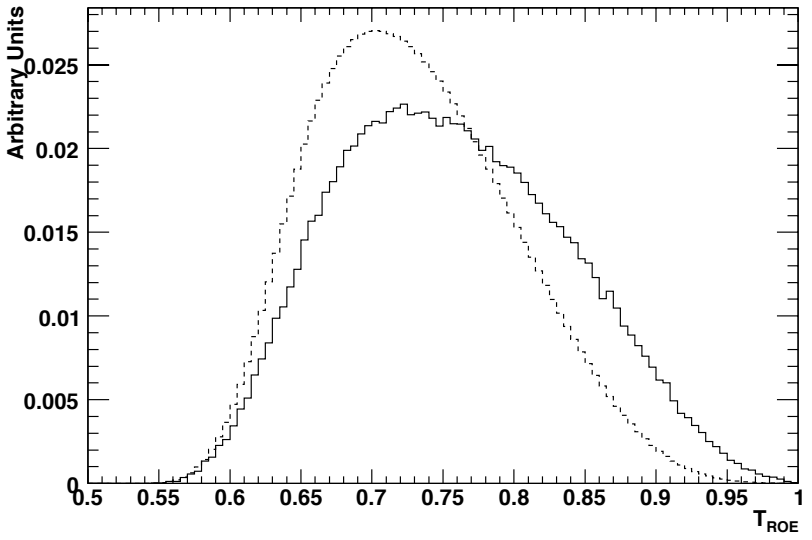


FIGURE 3.4: Normalised distributions of T_{ROE} for continuum (dashed line) and correctly reconstructed candidates (solid line) before post-reconstruction cuts.

The thrust of N detected particles each with momentum $\mathbf{p}_i$ is defined as,

$$T = \frac{\sum_i^N |\mathbf{n} \cdot \mathbf{p}_i|}{\sum_i^N |\mathbf{p}_i|}, \tag{3.4}$$

where $\mathbf{n}$ is the unit vector which maximises the value of T and whose direction defines the thrust axis of those particles [52]. Figure 3.4 shows the calculated thrust for the ROE, T_{ROE}. It can be seen that on average continuum events have a lower thrust for the ROE than correctly reconstructed signal events as they are less isotropic.

Another useful variable derived from the calculated thrust is the cosine of the angle between the photon momentum and the thrust axis of the ROE, $|\cos\theta_T|$. This has a uniform distribution for correctly reconstructed signal events whereas jet-like continuum events peak at unity. This difference is illustrated in figure 3.5.

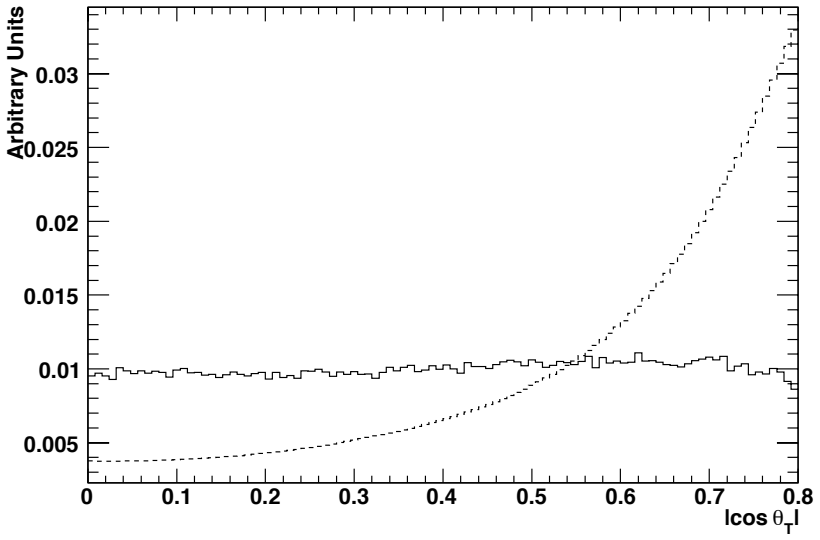


FIGURE 3.5: Normalised distributions of $|\cos\theta_T|$ for continuum (dashed line) and correctly reconstructed candidates (solid line) before post-reconstruction cuts.

3.3.2.2 Energy and Momentum Flow

A number of useful topological variables measure the angular distribution of energy and momentum flow in a given event. This analysis makes use of Fox-Wolfram moments [53] and both longitudinal and perpendicular 'monomial' functions [54].

For N particles with total energy E_{all}, the lth normalised Fox-Wolfram moment is defined as,

$$H_l = \sum_{i,j}^{N} \frac{|\mathbf{p}_i||\mathbf{p}_j|}{E_{all}} \mathcal{P}_l(\cos\theta_{ij}), \tag{3.5}$$

where $\mathcal{P}_l$ are Legendre polynomials of order l and θ_{ij} is the angle subtended between the momentum vectors of particles i and j. The ratio of the second to zeroth Fox-Wolfram moments, $R_2 = H_2/H_0$, is commonly used to measure the jet-like nature of an event. To reduce ISR backgrounds R_2 is calculated in the frame recoiling against the high energy photon for all particles excluding the photon. To distinguish this frame from the CM frame a prime notation is used, R_2'. Figure 3.6 shows the simulated distributions of R_2'. It can be seen that the signal distribution peaks closer to zero relative to the continuum background.

The normalised monomial functions are calculated relative to a given axis, α. For N particles whose momenta $\mathbf{p}_i$ subtend an angle θ_i with respect to α the lth monomial is

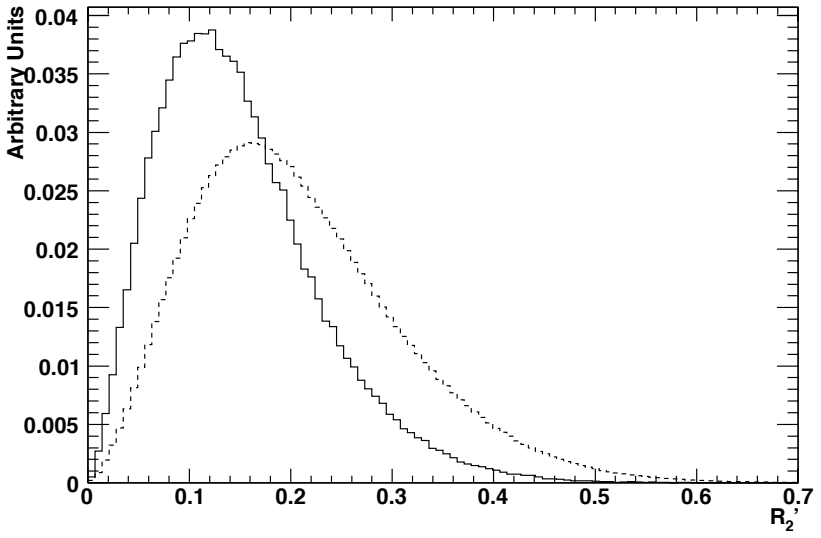


FIGURE 3.6: Normalised distributions of R_2' for continuum (dashed line) and correctly reconstructed candidates (solid line) before post-reconstruction cuts.

defined as,

$$L_l^\alpha = \frac{\sum_i^N |\mathbf{p}_i| \cos^l \theta_i}{\sum_i^N |\mathbf{p}_i|}, \tag{3.6}$$

$$P_l^\alpha = \frac{\sum_i^N |\mathbf{p}_i| \sin^l \theta_i}{\sum_i^N |\mathbf{p}_i|}, \tag{3.7}$$

for the longitudinal and perpendicular monomials respectively. This analysis considers the monomials functions of the ROE with respect to the thrust axis of the ROE and with respect to the high energy photon momentum axis. Figures comparing distributions for correctly reconstructed candidates and continuum backgrounds for each of the twelve monomials are contained in appendix A.

3.3.2.3 Sphericity Tensor

The sphericity tensor of an event is defined as,

$$S^{\alpha\beta} = \frac{\sum_i^N p_i^\alpha p_i^\beta}{\sum_i^N |\mathbf{p}_i|^2}, \tag{3.8}$$

where the indices α, β=1,2,3 denote the x, y and z three momentum components of the ith particle [55]. The resulting 3×3 matrix can be diagonalised giving three eigenvalues, $\lambda_1 \geq \lambda_2 \geq \lambda_3$ with $\lambda_1 + \lambda_2 + \lambda_3 = 1$. From these eigenvalues the sphericity, S, planarity,

P and aplanarity, A, for N particles are defined respectively as,

$$S = \frac{3}{2}(\lambda_1 + \lambda_2), \tag{3.9}$$

$$P = \lambda_2 - \lambda_1, \tag{3.10}$$

$$A = \frac{3}{2}\lambda_1. \tag{3.11}$$

Each of these variables has a distinct distribution for isotropic and jet-like events due to differences in the typical sphericity tensor for each topology. Figures showing the distributions of each sphericity variable for the ROE, comparing continuum and correctly reconstructed candidates, are contained in appendix A.

3.3.2.4 Angular Momentum

A final topological variable presents itself from angular momentum, the cosine of the angle between the signal candidate B meson momentum and the principal axis, $|\cos\theta_B|$. In a signal event the spin-1 $\Upsilon(4S)$ decays to two spin-0 B mesons hence the angular distributions of B mesons is proportional to $\sin^2\theta_B$, while fake B meson candidates from continuum decays are generally observed to have uniform distribution in $|\cos\theta_B|$. This distinction can be seen in figure 3.7 which shows simulated distributions of $|\cos\theta_B|$.

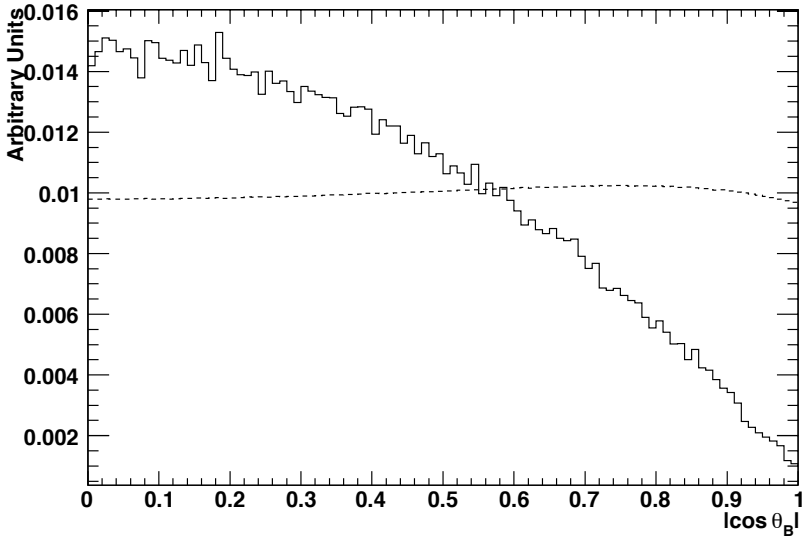


FIGURE 3.7: Normalised distributions of $|\cos\theta_B|$ for continuum (dashed line) and correctly reconstructed candidates (solid line) before post-reconstruction cuts.

3.3.3 Flavour Tagging Variables

Many analyses of B meson decays in *BABAR* data require the flavour of both B mesons to be known from their respective decays. Typically one B meson is fully reconstructed and knowledge of the flavour of the other B meson in an event is required. The properties of this particle, the so-called 'tag B', can be inferred from the ROE and a variety of tools exist in the *BABAR* framework to carry out this task. Although the analysis presented here does not explicitly depend on knowledge of the tag B, its properties can aid in the rejection of combinatoric backgrounds from continuum events.

Variables used to determine the flavour of the tag B meson aim to establish the b quark flavour by identifying correlations between the quark charge and the signed characteristics of the decay products [56]. The dedicated *BABAR* tagging group combine such variables into multivariate classifiers, known as sub-taggers, with different sub-taggers corresponding to different physics processes. The sub-taggers are subsequently combined to assign a probability to the flavour of the tag B meson. Importantly for this analysis, the sub-taggers have distinct distributions for events containing B meson decays compared to continuum events. Their outputs are thus useful discriminating variables. The complete tagging framework is described in detail in [57], an overview of seven sub-taggers used for this analysis is given here.

3.3.3.1 Lepton Tagging

Three of the sub-taggers rely on the lepton content of B meson decays. In $b \to c$ transitions a direct semi-leptonic decay will result in a lepton with the same sign charge as the parent b quark. Alternatively a lepton from a cascade $b \to c \to s$ decay can have either the same or opposite sign charge; however, it will exhibit a softer momentum spectrum. It is therefore possible to use the charge and kinematics of leptons from the ROE to identify the b quark flavour of the tag B meson. On average continuum events contain significantly fewer leptons, it follows that a classifier used to identify the lepton content of an event will have different performance on such data compared to signal B meson decay events.

An electron sub-tagger, L^e_{TAG}, has been trained to identify events where PID selectors identify any track in the ROE to be an electron or positron. The classifier is a neural net with input variables that give kinematic information of the assumed semi-leptonic decay. These variables, all calculated in the CM frame, are the momentum of the electron track candidate, the energy in the hemisphere defined by the direction of the virtual W boson in the assumed semi-leptonic decay and the cosine of the angle between the assumed

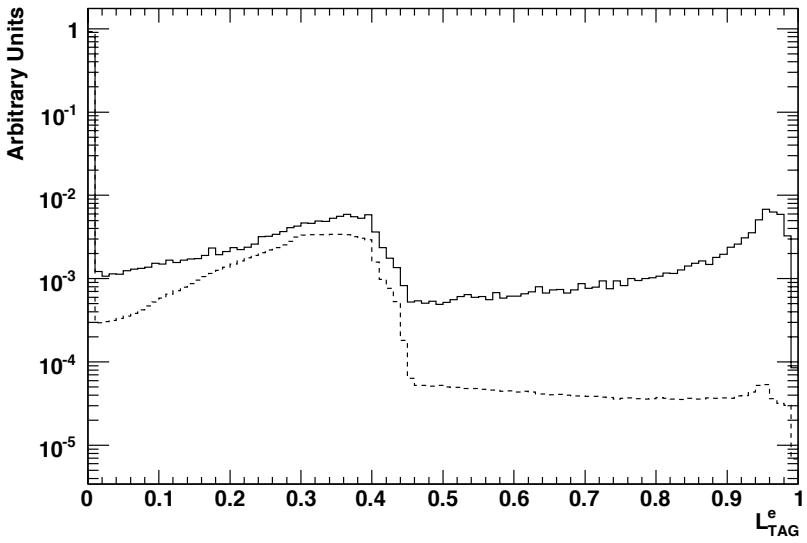


FIGURE 3.8: Normalised distributions of L^{e}_{TAG} for continuum (dashed line) and correctly reconstructed candidates (solid line) before post-reconstruction cuts. Note the log scale.

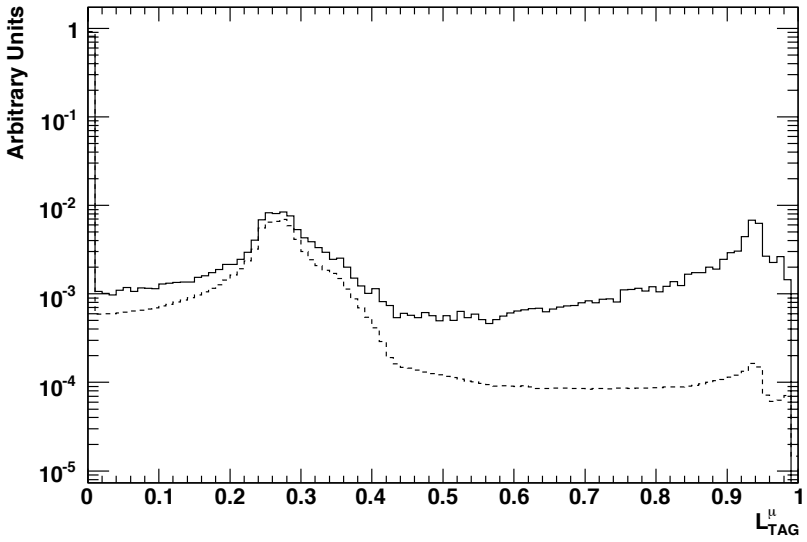


FIGURE 3.9: Normalised distributions of L^{μ}_{TAG} for continuum (dashed line) and correctly reconstructed candidates (solid line) before post-reconstruction cuts. Note the log scale.

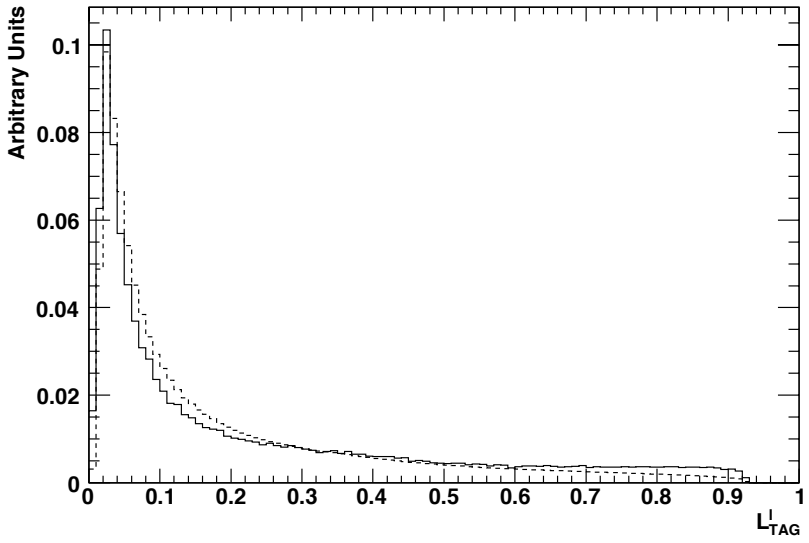


FIGURE 3.10: Normalised distributions of L^l_{TAG} for continuum (dashed line) and correctly reconstructed candidates (solid line) before post-reconstruction cuts.

electron track momentum and the event missing momentum. Figure 3.8 shows the output of this classifier for simulated data and compares the distributions for correctly reconstructed candidates and continuum backgrounds. In events where the tag B meson is correctly identified the output peaks at one for direct decays with a secondary peak at ~ 0.35 due to cascade decays. The large peak at zero represents events where no electron or positron were identified. The output is typically multiplied by the charge of the track to separate electron and positrons when identifying flavour although this is not shown here. It can be seen that there is an order of magnitude difference in the sub-tagger giving a positive identification of direct leptonic decays between these data. The presence of the secondary peak in the continuum distribution is mainly a consequence of semi-leptonic D meson decays in these data. A similar sub-tagger, L^{μ}_{TAG}, has been trained for events where any track in the ROE is identified as a muon. Figure 3.9 shows the output of the muon sub-tagger for the same data. The structure of these distributions distinguishes muons from direct and cascade decays in the same way as the electron sub-tagger. Again a significant difference in positive classifier output is observed. A final lepton sub-tagger, L^l_{TAG}, is trained using the same kinematic variables for tracks not identified as leptons in an attempt to recover any flavour information lost due to PID inefficiencies. The most lepton-like track is selected by this classifier, the output of which is shown for simulated data in figure 3.10. A value closer to one is observed for leptons which have been correctly recovered. For this sub-tagger the discrimination between signal and continuum data is not as significant; however, a residual difference at higher output values is still observed.

3.3.3.2 Hadron Tagging

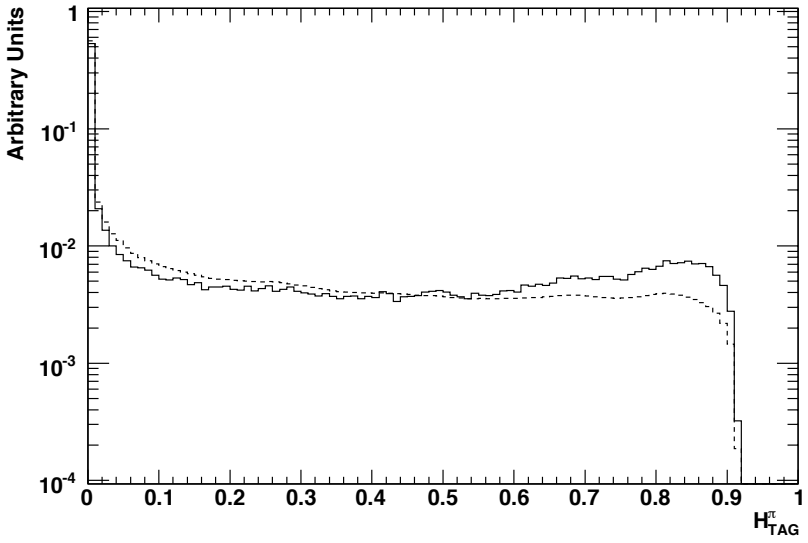


FIGURE 3.11: Normalised distributions of H^{π}_{TAG} for continuum (dashed line) and correctly reconstructed candidates (solid line) before post-reconstruction cuts. Note the log scale.

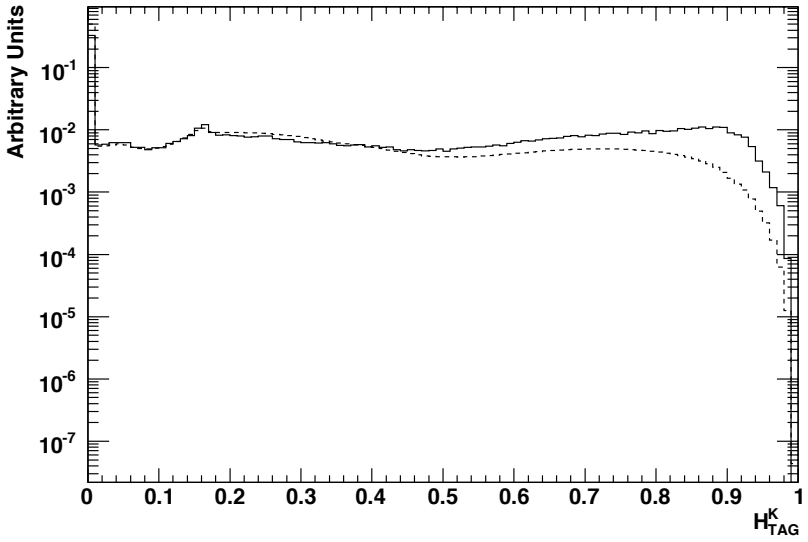


FIGURE 3.12: Normalised distributions of H^{K}_{TAG} for continuum (dashed line) and correctly reconstructed candidates (solid line) before post-reconstruction cuts. Note the log scale.

Decays of B mesons containing a $D^{*\pm}$ will often have an associated low momentum charged pion from the subsequent $D^{*\pm}$ decay. This pion will always have opposite sign charge to the parent b quark. The slow pion sub-tagger, H^{π}_{TAG}, makes use of this physics

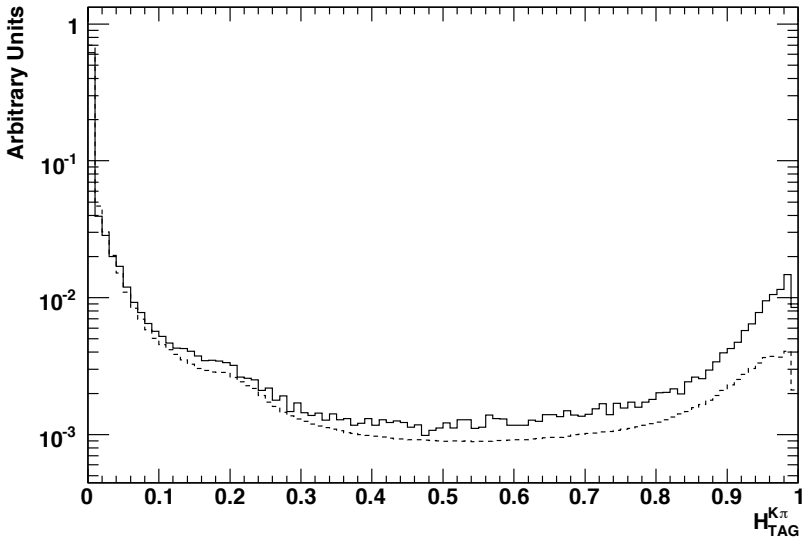


FIGURE 3.13: Normalised distributions of $H^{K\pi}_{TAG}$ for continuum (dashed line) and correctly reconstructed candidates (solid line) before post-reconstruction cuts. Note the log scale.

signature by identifying tracks of low CM momentum correlated with the thrust axis of all other tracks and neutrals from the ROE in that frame. Two variables describing these kinematics, along with a kaon PID classifier output to reject low momentum charged kaons, are combined in a neural net, the output of which is shown in figure 3.11 for correctly reconstructed and continuum candidates. The peak at 0.9 is observed for candidates where a low momentum track was consistent with the physics hypothesis, while the peak at zero is due to candidates with either no suitable low momentum track in the ROE or where the low momentum track identified was not consistent with a pion from a $D^{*\pm}$ decay. A similar kaon sub-tagger, H^{K}_{TAG}, identifies charged kaons from $b \to c \to s$ transitions which will have the same charge sign as the parent b quark. This tagger combines kaon PID information and the transverse momentum properties of candidate tracks. The kaon sub-tagger distribution is shown in figure 3.12; for signal decays it peaks near one when kaons from such transitions are identified in the ROE. Outputs of the kaon and slow pion sub-taggers along with information of the angular separation of the identified pions and kaons are combined into a neural net to form a 'kaon slow pion' sub-tagger, $H^{K\pi}_{TAG}$. This sub-tagger aims to take advantage of the correlation between charge sign expected for both input sub-taggers for a true $D^{*\pm}$ decay from the tag B meson. The output of the kaon slow pion sub-tagger is shown in figure 3.13. The peak at one for signal indicates a correctly identified tag B meson. For each hadronic sub-tagger there is on average an excess in positive output from signal decays compared to continuum events; however, this is not as significant as the difference observed in the

electron and muon sub-tagger outputs. This is because continuum events can contain pions and kaons with similar kinematics to the $D^{*\pm}$ daughters in a B meson decay chain; however, their angular distributions relative to the thrust axis are typically different due to the continuum jet-like nature resulting in the limited discrimination observed.

3.3.3.3 Momentum Tagging

The final sub-tagger, P_{TAG}, is used to identify high momentum particles, such as high momentum pions from $B \to D^*\pi$ decays. High energy tracks typically originate from the primary B meson decay and tend to have the same sign charge as the parent b quark. They will also have a small transverse impact parameter due to their prompt production and their direction will be correlated with that of the tag B meson as they carry away a large proportion of its momentum. The high momentum sub-tagger is a neural net with three inputs reflecting these expected physical properties. The sub-tagger output is shown in figure 3.14, a higher output indicates the presence of a high momentum track in the ROE while the absence of such a particle is reflected in the peak at zero. For continuum events it can be seen that the output is typically less than for signal events and the discriminatory power between these event types is similar to that of the hadronic sub-taggers.

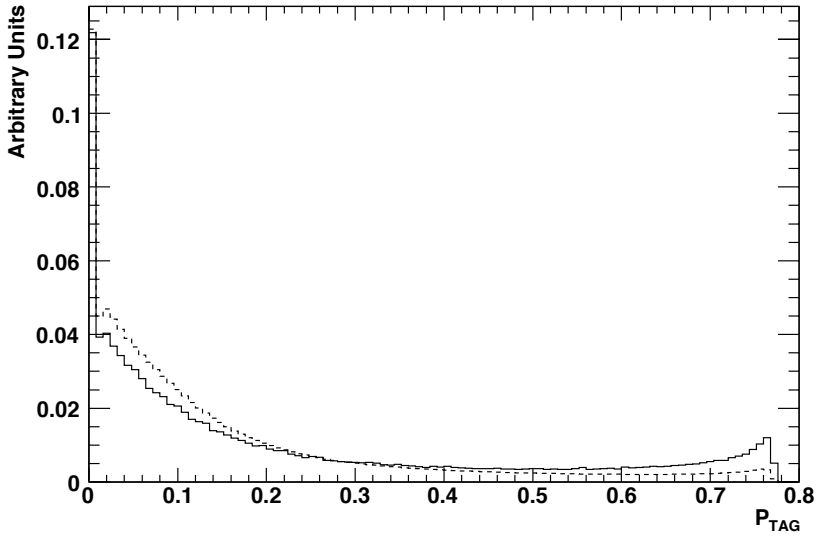


FIGURE 3.14: Normalised distributions of P_{TAG} for continuum (dashed line) and correctly reconstructed candidates (solid line) before post-reconstruction cuts.

3.4 Event Reduction Filters

As a consequence of computer processing and storage limitations, pre-reconstruction event filters[4] must be applied to both experimental and MC data before full event reconstruction. General event filters are called from the *BABAR* framework. The `BGFMultiHadron` [58] passes events which have at least three charged tracks and $R_2 < 0.98$ for all charged tracks, essentially eliminating Bhabha scattering backgrounds. Events are then tested for consistency with criteria required for B counting [45] through the filter `BCountMHFilter` [58]. A custom filter developed for previous *BABAR* analyses studying $B \to \rho\gamma$ transitions, `BtoRhoGammaBToXGF` [59], requires that for all particles in an event $R_2 < 0.9$ and that the highest energy neutral cluster in the EMC has CM energy in the range $1.15 < E^* < 3.5\,\text{GeV}$. All events passing these filters are then passed through a custom filter developed for this analysis, `BToXdGammaFilter`.

3.4.1 `BToXdGammaFilter`

The filter `BToXdGammaFilter` is optimised to efficiently reject backgrounds and accept simulated data events without exceeding centrally assigned processing constraints for both MC and experimental data. In each event B meson candidates are reconstructed for all $B \to X_d\gamma$ modes listed in table 3.1. The hadronic X_d components are reconstructed combining tracks from `GoodTracksLoose`, π^0 candidates from `pi0DefaultMass` and η candidates from `etaggDefault` with all candidates from these lists required to have momentum greater than $0.3\,\text{GeV}/c$ in the lab frame. Where a mode has more than one charged track, X_d candidate decay vertices are fitted with the `Cascade` algorithm applying the geometric constraint and requiring the χ^2 probability of the fit to be greater than 0.001. Where there is only one track, X_d candidates are constructed by adding the four-momentum of the daughter particles. The composite X_d candidates are required to have an invariant mass less that $2.2\,\text{GeV}/c^2$.

Reconstructed X_d candidates are combined with photons to form B meson candidates. Photons, from `GoodPhotonLoose`, are required to have CM energy in the range $1.15 < E^* < 3.5\,\text{GeV}$ and clusters with at least four EMC crystals over energy readout threshold to eliminate fake neutral clusters which can potentially arise due to a single noisy crystal. B meson candidates are required to have $|\Delta E| < 0.3\,\text{GeV}$ and $5.0 < m_{ES} < 5.3\,\text{GeV}/c^2$. The angle between the B meson candidate momentum and thrust axis of the ROE is required to have an absolute cosine less than 0.8. Any event which has at least one

[4] Pre-reconstruction (post-reconstruction) refers to event processing carried out before (after) the custom event reconstruction described in section 3.5. All events discussed in this chapter have been reconstructed within the central *BABAR* framework.

MC data	Filter	Events	Relative Eff (%)	Total Eff (%)
$B^+ \to X_{du}\gamma$	None	1,288,000	100.0	100.0
	`BGFMultiHadron`	1,244,491	96.6	96.6
	`BCountMHFilter`	1,149,425	92.4	89.2
	`BtoRhoGammaBToXGF`	959,714	83.5	74.5
	`BToXdGammaFilter`	453,992	47.3	35.2
$B^0 \to X_{dd}\gamma$	None	1,288,000	100.0	100.0
	`BGFMultiHadron`	1,241,756	96.4	96.4
	`BCountMHFilter`	1,143,968	92.2	88.9
	`BtoRhoGammaBToXGF`	951,564	83.2	73.9
	`BToXdGammaFilter`	441,343	46.4	34.3
$B^+ \to X_{su}\gamma$	None	11,178,000	100.0	100.0
	`BGFMultiHadron`	10,762,617	96.3	96.3
	`BCountMHFilter`	9,804,116	91.1	87.7
	`BtoRhoGammaBToXGF`	8,136,455	83.0	72.8
	`BToXdGammaFilter`	3,293,162	40.5	29.5
$B^0 \to X_{sd}\gamma$	None	11,178,000	100.0	100.0
	`BGFMultiHadron`	10,724,379	95.9	95.9
	`BCountMHFilter`	9,726,955	90.7	87.0
	`BtoRhoGammaBToXGF`	8,060,449	82.9	72.1
	`BToXdGammaFilter`	3,279,972	40.7	29.3

TABLE 3.5: Efficiencies of pre-reconstruction event filters when applied to the high mass bin signal MC data.

B meson candidate passing all of the above filter requirements is kept for full event reconstruction.

The choice of variables used in `BToXdGammaFilter` and the value of cuts on those variables are based primarily on the equivalent filter used in the previous version of this analysis [1]. However, the filter described here improves on the previous analysis through a revision of the choice of *BABAR* event framework track and neutral candidate lists used for hadronic candidate reconstruction. The lists described above perform tighter selections on their constituent candidates and are therefore smaller in size. This leads to a significant reduction in the potential number of combinations allowed to form hadronic candidates in a given event with an almost negligible cost in signal efficiency. The processing overhead saved allows the invariant mass cut to be relaxed from $2.0\,\mathrm{GeV}/c^2$ to $2.2\,\mathrm{GeV}/c^2$ therefore allowing for a potentially more inclusive measurement. Due to the required processing constraints this cut cannot be relaxed any further.

MC data	Filter	Events	Relative Eff (%)	Total Eff (%)
$B^0 \to \rho^+\gamma$	None	650,000	100.0	100.0
	`BGFMultiHadron`	609,153	93.7	93.7
	`BCountMHFilter`	508,536	83.5	78.2
	`BtoRhoGammaBToXGF`	443,283	87.2	68.2
	`BToXdGammaFilter`	323,325	72.9	49.7
$B^0 \to \rho^0\gamma$	None	650,000	100.0	100.0
	`BGFMultiHadron`	636,882	98.0	98.0
	`BCountMHFilter`	539,034	84.6	82.9
	`BtoRhoGammaBToXGF`	450,721	83.6	69.3
	`BToXdGammaFilter`	364,193	80.8	56.0
$B^0 \to \omega\gamma$	None	650,000	100.0	100.0
	`BGFMultiHadron`	622,903	95.8	95.8
	`BCountMHFilter`	513,547	82.4	79.0
	`BtoRhoGammaBToXGF`	439,831	85.6	67.7
	`BToXdGammaFilter`	318,105	72.3	48.9
$B^0 \to K^{*+}\gamma$	None	6,449,000	100.0	100.0
	`BGFMultiHadron`	6,082,909	94.3	94.3
	`BCountMHFilter`	5,059,206	83.2	78.4
	`BtoRhoGammaBToXGF`	4,277,607	84.6	66.3
	`BToXdGammaFilter`	2,136,768	50.0	33.1
$B^0 \to K^{*0}\gamma$	None	6,449,000	100.0	100.0
	`BGFMultiHadron`	6,047,163	93.8	93.8
	`BCountMHFilter`	4,938,077	81.7	76.6
	`BtoRhoGammaBToXGF`	4,149,053	84.0	64.3
	`BToXdGammaFilter`	2,764,181	66.6	42.9

TABLE 3.6: Efficiencies of pre-reconstruction event filters when applied to the low mass bin signal MC data.

MC data	Events	Eff (%)
$B^+ \to X_{du}\gamma$	412,189	34.3
$B^0 \to X_{dd}\gamma$	398,935	33.2
$B^+ \to X_{su}\gamma$	3,042,628	29.2
$B^0 \to X_{sd}\gamma$	2,983,045	28.6

TABLE 3.7: High mass bin signal MC samples after filter and hadronic mass cuts. The efficiency is relative to the pre-filter samples given in table 3.3.

MC data	Events	Eff (%)
$B^0 \to \rho^+\gamma$	282,157	49.6
$B^0 \to \rho^0\gamma$	317,817	55.8
$B^0 \to \omega\gamma$	318,105	48.9
$B^0 \to K^{*+}\gamma$	1,889,573	32.8
$B^0 \to K^{*0}\gamma$	2,441,107	42.6

TABLE 3.8: Low mass bin signal MC samples after filter and hadronic mass cuts. The efficiency is relative to the pre-filter samples given in table 3.4.

Type	Pre Filter Events	Post Filter Events	Efficiency (%)
Generic B^+B^- MC	708,762,000	1,623,556	0.23
Generic $B^0\overline{B}^0$ MC	717,995,000	1,081,134	0.15
Generic $c\bar{c}$ MC	1,128,544,000	9,642,971	0.85
Generic $q\overline{q}$ MC ($q = u, d, s$)	1,670,948,000	27,897,989	1.68
On Peak Data	6,714,057,036	20,033,960	0.30
Off Peak Data	642,952,515	1,932,935	0.30

TABLE 3.9: Pre-reconstruction filter efficiencies for generic MC backgrounds and experimental data.

3.4.2 Filter Efficiencies

The relative and cumulative efficiencies of the above filters when applied to the signal MC data listed in table 3.2 are shown in tables 3.5 and 3.6 for high mass bin and low mass bin samples respectively. After the filters are applied the model cut on hadronic mass due to m_X^T=1.0 GeV/c^2 must be made to the filtered samples. Tables 3.7 and 3.8 show the effect of this cut for the high and low mass bin samples respectively; the efficiencies quoted in these tables are the effective combined filter efficiency of each sample for events which satisfy the mass threshold requirement. Table 3.9 shows the total combined efficiency of all filters when applied to the generic MC data samples listed in table 3.2 and to the experimental data.

The combined signal efficiencies appear low due to low efficiency for events where the signal decay does not directly correspond to one of the reconstructed modes listed in table 3.1.

3.5 Full Event Reconstruction

Full event reconstruction is performed for all events passing the pre-reconstruction filters. Hadronic candidates are created with identical lists and using the same vertex fitting

strategy as the `BToXdGammaFilter` event filter. However, the laboratory momentum cut is relaxed to $0.1\,\mathrm{GeV}/c$, the vertex χ^2 probability cut is not applied and candidates can have a hadronic mass up to $2.4\,\mathrm{GeV}/c^2$. The relaxation of these cuts is temporary and a historical artifact of the code inherited from the previous version of this analysis [1]; they are tightened again to reflect the filter cuts in the subsequent post-reconstruction event selection. Additionally where there was only one charged final state particle the candidates are formed by adding the daughter four-momenta constraining all daughters to originate from the event primary vertex.

Before hadronic candidates are fitted a PID selection is applied to identify them as either X_d or X_s candidates. For each track in `GoodTracksLoose` considered for hadronic candidate reconstruction the PID selectors `pionKMLoose` and `kaonBDTNotAPion` are called. If a track passes both or neither selectors it is not considered for reconstruction. Tracks passing only the pion selector are assumed to be pions. Tracks passing only the kaon selector are assumed to be kaons and their four-momentum is redefined by assigning them the world average kaon mass [14]. During hadronic candidate reconstruction if all tracks are identified as pions the candidate is classified as a X_d and fitted. If all but one track is a pion and the remaining track a kaon with correct sign charge where relevant, the candidate is classified as a X_s and fitted. Any other combination of tracks results in the candidate being vetoed from the reconstruction. Fitted hadronic candidates satisfying the reconstruction criteria are combined with photons from `GoodPhotonLoose` applying the same CM frame energy cut on the photon as in `BToXdGammaFilter` to form B meson candidates. Again the addition of these four momenta are constrained such that the X_d candidate and photon candidate originate from the event primary vertex. B meson candidates are required to have m'_{ES} greater than $5.0\,\mathrm{GeV}/c^2$ and $|\Delta E| \leq 0.3\,\mathrm{GeV}$. A further cut of $|\cos\theta_T| \leq 0.8$ is then applied to reduce continuum backgrounds in order to save disk space and reduce processing time in subsequent event selection. This cut does not exactly mirror the similar cut made in `BToXdGammaFilter` as here the angle is defined as the thrust axis relative to the photon direction as opposed to the reconstructed B meson direction; however, these distributions are similar as the background photon direction is generally correlated with jet momentum in continuum events whereas signal photon momentum direction will on average reflect the isotropic nature of signal events. For each reconstructed B meson candidate satisfying the above requirements kinematic, topological and tagging variable information is saved for subsequent analysis. Table 3.10 shows the total number of events containing at least one B meson candidate after reconstruction for both MC and experimental data.

Data Type	≥ 1 Cand	Eff (%)	≥ 1 X_d Cand	≥ 1 X_s Cand
$X_{du}\gamma$ signal MC	334,022	81.0	287,813 (90,200)	152,148 (2,476)
$X_{dd}\gamma$ signal MC	325,232	81.5	287,625 (60,822)	134,698 (3,016)
$X_{su}\gamma$ signal MC	2,374,495	78.0	1,275,415 (47,203)	1,808,478 (584,981)
$X_{sd}\gamma$ signal MC	2,341,138	78.5	1,256,499 (29,861)	1,772,401 (486,733)
$\rho^+\gamma$ signal MC	249,863	88.6	237,867 (147,054)	90,279 (3,957)
$\rho^0\gamma$ signal MC	272,947	85.9	258,797 (197,827)	80,571 (10,809)
$\omega\gamma$ signal MC	273,002	85.8	258,090 (46,600)	101,574 (1,552)
$K^{*+}\gamma$ signal MC	1,574,018	83.3	798,196 (109,072)	1,135,465 (466,752)
$K^{*0}\gamma$ signal MC	2,044,754	83.8	644,443 (30,181)	1,806,368 (1,323,321)
Generic B^+B^- MC	1,044,389	64.3	553,997	697,965
Generic $B^0\overline{B}^0$ MC	709,728	65.6	390,398	463,011
Generic $c\bar{c}$ MC	6,490,155	67.3	4,085,150	3,799,608
Generic $q\bar{q}$ MC	19,001,795	68.1	15,163,515	6,284,873
On Peak Data	13,702,098	68.4	9,976,160	5,790,300
Off Peak Data	1,323,267	68.5	974,466	541,629

TABLE 3.10: Events with at least one reconstructed candidate after full event reconstruction for both MC and experimental data and the reconstruction efficiency relative to the number of events passing pre-reconstruction event filters given in tables 3.7, 3.8 and 3.9. Also shown are the number of events with at least one fitted X_d candidate and the number of events with at least one fitted X_s candidate. Shown in parentheses are the number of events where a correctly reconstructed candidate is identified. Note that the PID is not required to be correct for a candidate to be considered correctly reconstructed.

3.5.1 Signal Candidate Identification

When considering signal MC data the identification of correctly reconstructed events is vital to estimate the combinatoric background contributions from reconstructing and selecting candidates within these data. Due to contamination from beam backgrounds and detector noise there is no one-to-one correspondence between particles generated in underlying physics processes and those in the reconstructed lists available after the detector simulation has run. For this reason the *BABAR* simulation framework uses objects known as '`gHits`' for each simulated particle to model its interaction with active detector elements. Detector responses such as DCH hits and EMC crystal readouts can have one, many or no `gHits` associated with them, depending on the origin of the interaction and the simulated efficiencies of that particular subsystem. The number of `gHits` present in reconstructed tracks and clusters allow the user to find consistency between that reconstructed candidate and any generated particle.

This analysis has two dominant combinatoric backgrounds in signal MC data. Firstly a candidate whose reconstructed mode matches that of the generated event but which

uses one or more tracks and clusters from the ROE in the fit. Secondly a candidate whose reconstructed mode does not match the generated mode. Of the latter there are those candidates where the generated mode was one of the other signal modes from table 3.1 and those where the generated mode was some other allowed $B \to X_{s/d}\gamma$ transition. No significant contribution is observed from candidates where the hadronic component is correctly reconstructed but an incorrect photon is used to form the B meson. Combinatoric backgrounds are therefore dominated by the reconstruction of the hadronic candidate.

During the reconstruction of each signal event a list of the generated particles is used to identify the true decay mode. If the generated mode is a signal mode and there is a reconstructed candidate of the same mode then the `gHit` consistency associator is used to identify if the hadronic component was correctly reconstructed. The associator is called for each reconstructed track and for each cluster used to create a π^0 or η. If all reconstructed candidates are associated with a unique generated charged pion, charged kaon or photon from the true signal B meson then the hadronic component and its corresponding B meson are considered to be correctly reconstructed.

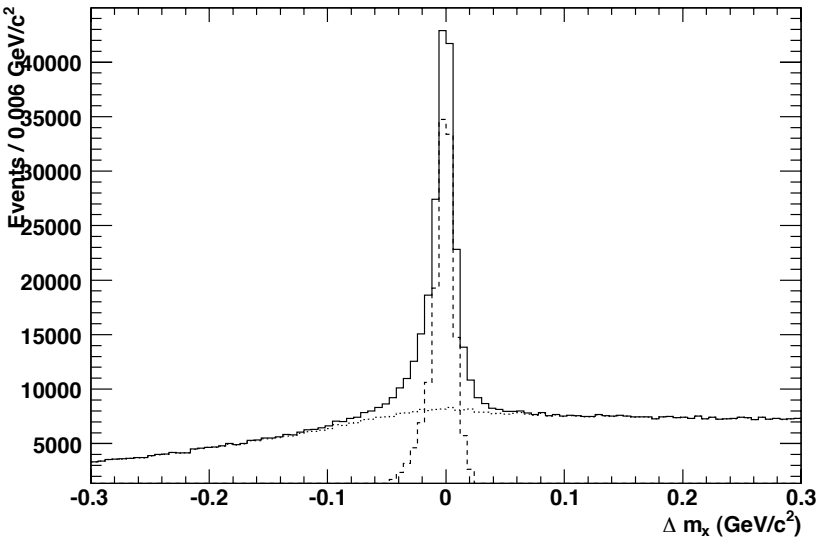


FIGURE 3.15: Distribution of Δm_X for X_d candidates whose reconstructed mode matches the generated mode in high mass $B \to X_d\gamma$ simulated events. The dashed line shows candidates identified as correctly reconstructed by `gHit` association and the dotted line shows the contribution from all other candidates where the reconstructed decay mode matches the generated decay mode.

To investigate the `gHit` associator method for inefficiencies an independent cross check is carried out. The difference between the reconstructed and generated hadronic mass, Δm_X, for candidates whose reconstructed mode matches the generated mode of an

event will peak at zero with a narrow spread due to detector resolution for correctly reconstructed candidates, while no significant narrow peak will exist for background candidates of that mode. Figure 3.15 shows the distribution of Δm_X for $B \to X_d\gamma$ signal MC candidates where the reconstructed mode matches the generated mode. It can be seen that candidates that fail the `gHit` associator do not show a significant peak at zero. It is therefore assumed that for the purposes of this analysis the associator efficiently identifies correctly reconstructed candidates.

3.6 Post Reconstruction Event Reduction

Combinatoric backgrounds dominate over events with correctly reconstructed candidates after full event reconstruction. In order to maximise the statistical significance of correctly reconstructed events a number of variables are identified to which cuts can be applied. The following section describes the variables used and details how the optimal values for each cut were determined.

3.6.1 $B \to X_s\gamma$ Background from K^0_S Decays

In both the high and low mass region, $B \to X_d\gamma$ candidate decays for certain modes will have hadronic final states identical to $B \to X_s\gamma$ decays where there was an intermediate $K^0_S \to \pi^+\pi^-$ or $K^0_S \to \pi^0\pi^0$ decay. For example the decay $B^0 \to X_d\gamma \to \pi^+\pi^-\pi^0\gamma$ is identical to the decay $B^0 \to X_s\gamma \to K^0_S(\pi^+\pi^-)\pi^0\gamma$. Such backgrounds can contribute in all reconstructed $B \to X_d\gamma$ modes with three or more final state particles from the hadronic decay.

The background from $K^0_S \to \pi^+\pi^-$ decays can be reduced by considering the decay length and invariant mass of potential K^0_S candidates. During event reconstruction all X_d candidates with two or more charged pions have their vertex reconstructed with the `Cascade` algorithm which determines the most probable position for the decay vertex of that candidate. Figure 3.16 shows the distance between the decay vertex and event primary vertex in the laboratory frame, x_{DEC}, for both correctly reconstructed $B \to X_d\gamma$ candidates and correctly reconstructed $B \to X_s\gamma$ decays whose final states are indistinguishable due an intermediate $K^0_S \to \pi^+\pi^-$ decay. It can be seen that the K^0_S backgrounds have a significant tail at high values of x_{DEC}, not seen in signal events. This can be attributed to the relatively large decay length of the K^0_S meson [14]. During event reconstruction pairs of oppositely charged tracks from `GoodTracksLoose` which pass the PID selector `PionKMLoose` (and therefore which can potentially be used in X_d candidate reconstruction) are also separately vertexed with the `Cascade` algorithm and

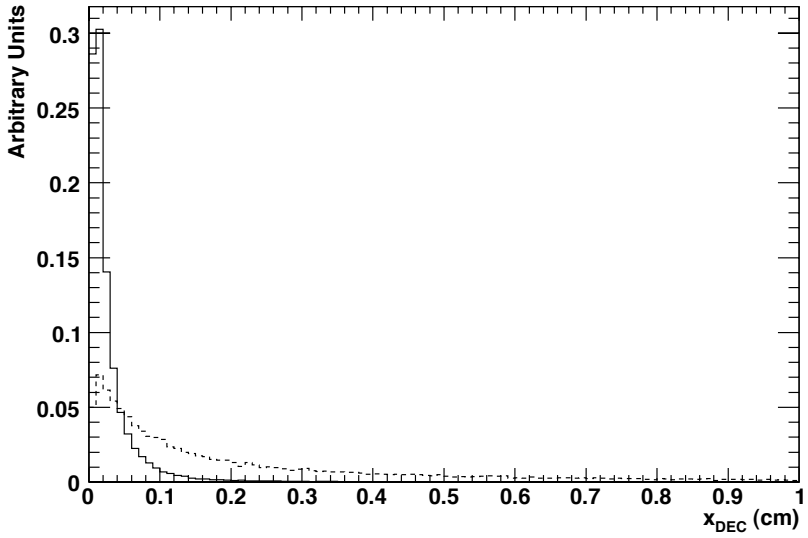


FIGURE 3.16: Distance between the decay vertex and event primary vertex for reconstructed modes with at least two charged tracks and at least three hadronic final state particles showing correctly reconstructed $B \to X_d\gamma$ candidates (solid line) and correctly reconstructed $B \to X_s\gamma$ background candidates which have indistinguishable final states due to an intermediate $K^0_S \to \pi^+\pi^-$ decay (dashed line).

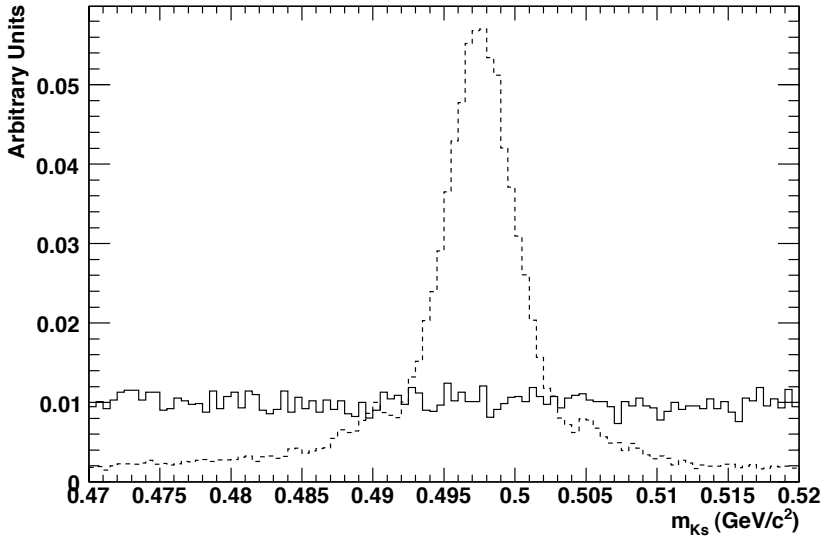


FIGURE 3.17: Invariant mass of reconstructed $K^0_S \to \pi^+\pi^-$ candidates where both charged tracks are also used to reconstruct a correctly identified $B \to X_d\gamma$ candidate (solid line) or a correctly identified $B \to X_s\gamma$ background candidate with an intermediate $K^0_S \to \pi^+\pi^-$ decay (dashed line)

saved. Figure 3.17 shows the the invariant mass of such track pairs, m_{K_s}, when both tracks are also used in the reconstruction of a signal $B \to X_d\gamma$ or the reconstruction of a correctly identified $B \to X_s\gamma$ background with intermediate $K^0_S \to \pi^+\pi^-$ decay. The background shows a clear peak about the K^0_S mass whereas those candidates where the pion pair came from a signal decay have a uniform distribution. Requiring that no X_d candidate with greater than three hadronic final state particles (of which at least two were charged pion candidates) is reconstructed from any pair of charged tracks which were also used to reconstruct a K^0_S candidate with $0.485 \leq m_{K_s} \leq 0.51\,\mathrm{GeV}/c^2$ and that the X_d candidate has $x_{DEC} \leq 0.2\,\mathrm{cm}$ is found to reject >95% of $K^0_S \to \pi^+\pi^-$ backgrounds while rejecting <5% of signal candidates whose modes are subject to this background. For ease of technical implementation the cut on x_{DEC} is also applied to mode 1 of table 3.1 at the cost of <5% of signal candidates from that mode before further post reconstruction cuts are applied.

Backgrounds from $K^0_S \to \pi^0\pi^0$ are less straight forward to reduce. This background only contains neutral particles so it is not possible to obtain reliable vertex information for the K^0_S decay. Consequently, as π^0 candidates in `Pi0DefaultMass` are defined to originate from the primary vertex, the mass resolution of any reconstructed K^0_S candidate will be much worse than those reconstructed from charged track pairs. Time constraints on the analysis meant no study was carried out to reduce this background, the consequences of which are discussed in section 3.6.7.1.

3.6.2 High Energy Photon Candidate

The dominant background for high energy photon candidates are asymmetric $\pi^0 \to \gamma\gamma$ and $\eta \to \gamma\gamma$ processes where one photon carries away most of the energy from the decay. To reduce these backgrounds every high energy photon candidate used to reconstruct a B meson has its four-momentum added to all other photons in the event with lab energy greater than 30 MeV (250 MeV) and the invariant mass closest to the nominal π^0 (η) mass [14] is saved. Figures 3.18 and 3.19 show the π^0 and η distributions respectively for combinatoric and correctly reconstructed MC events before post-reconstruction cuts are applied. During cut optimisation an invariant mass veto window is considered with each veto consisting of an upper and lower cut.

The lateral moment cut required for a cluster to be included in the `GoodPhotonLoose` (see section 3.2.2) list is tightened so that only photons with lateral moment less than 0.6 are considered. Other radiative penguin analyses using *BABAR* data have shown that this cut value has almost negligible signal rejection [59] and so it is implemented here as well.

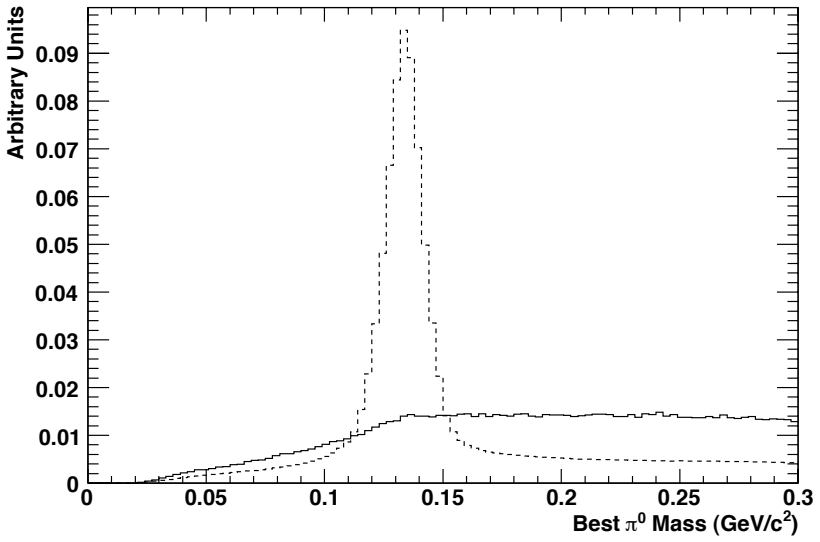


FIGURE 3.18: Normalised distributions of the invariant mass closest to the π^0 mass from combining the high energy photon with other photons of energy >30 MeV for continuum (dashed line) and correctly reconstructed candidates (solid line) before post-reconstruction cuts.

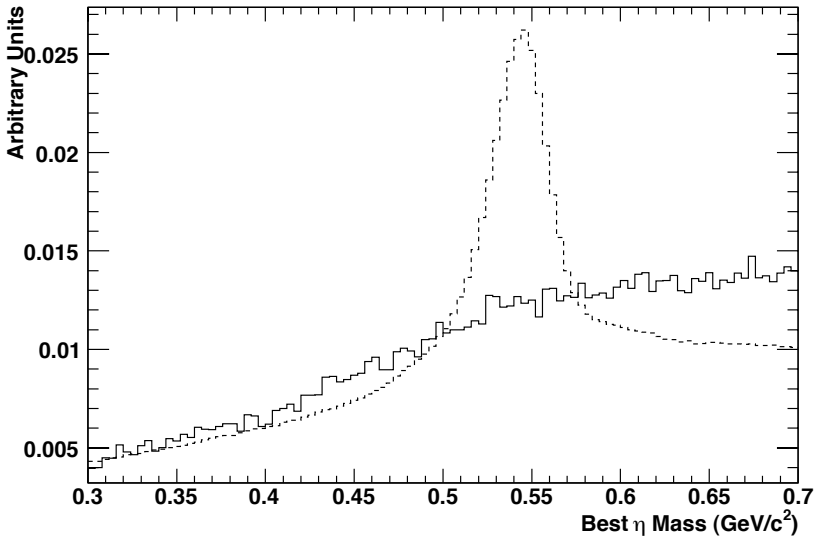


FIGURE 3.19: Normalised distributions of the invariant mass closest to the η mass from combining the high energy photon with other photons of energy >250 MeV for continuum (dashed line) and correctly reconstructed candidates (solid line) before post-reconstruction cuts.

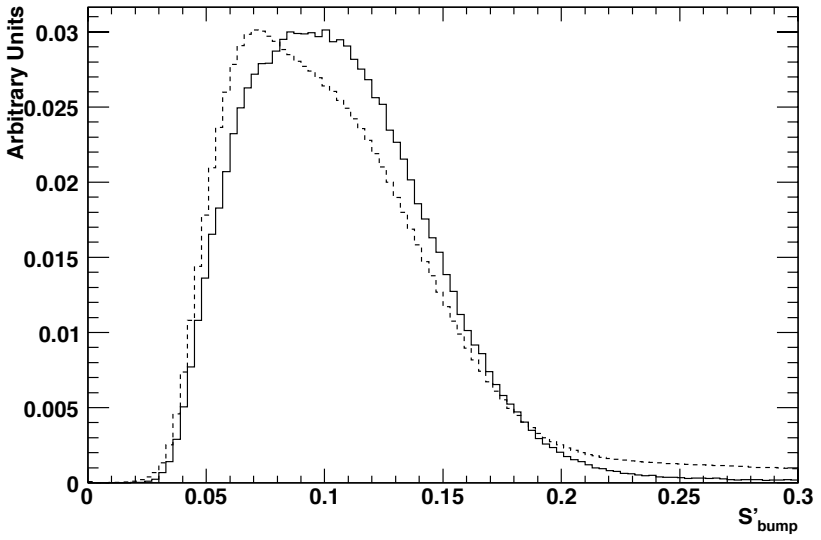


FIGURE 3.20: Normalised distributions of S'_{bump} for continuum (dashed line) and correctly reconstructed candidates (solid line) before post reconstruction cuts.

Backgrounds from merged $\pi^0 \to \gamma\gamma$ decays where both photons form the same cluster in the EMC can be rejected by considering the second moment of the cluster, S_{bump}. From equation (2.5) it can be seen that S_{bump} is geometrically dependent on the position of a cluster in the EMC. It is therefore more useful to multiply S_{bump} by the distance from the IR to the cluster centroid, giving a corrected second moment, S'_{bump}, independent of the geometry. Figure 3.20 shows the distributions of S'_{bump} for combinatoric and correctly reconstructed candidates before cuts are applied and it is this value for which a cut is considered. The enhanced high end tail in the combinatoric distribution can be mainly attributed to merged π^0 decays.

A further background which can contribute fake high energy photons arises from interactions between high energy charged particles and detector material whereby scattered neutrons can deposit energy in the EMC. If that cluster is then mistakenly matched with the charged track, the actual energy deposit from the charged particle will be assigned as a neutral particle. To reduce these backgrounds the distance from the candidate photon cluster to the nearest cluster associated with a charged track, x_{trk}, is a useful discriminant. An isolation cut requiring $x_{trk} \geq 25\,\text{cm}$ is applied based on the equivalent cut used in previous radiative penguin analyses using *BABAR* data [1, 59].

To remove possible backgrounds from single noisy crystals in data the EMC cluster is required to have greater than four constituent crystals. Additionally there must be no dead or damaged crystals within the cluster to ensure the energy measurement is as accurate as possible.

Finally ISR background photons typically have low transverse momentum so are largely found in the endcap. Therefore requiring that reconstructed photons only have associated clusters whose centroid is in the barrel of the EMC can reduce these backgrounds. Rejection of endcap photons is thus a binary variable and is considered in the cut optimisation discussed below.

3.6.3 $X_{s/d}$ Candidate Cuts

To further reduce combinatoric backgrounds cuts can be placed on reconstructed hadronic $X_{s/d}$ candidates. For those modes with more than one charged track a cut is placed on the χ^2 probability of the vertex fit. The cut value is based on the experience of previous radiative penguin analyses where multiple track final states are reconstructed [1, 30] and is required to be greater than 1% for all such candidates.

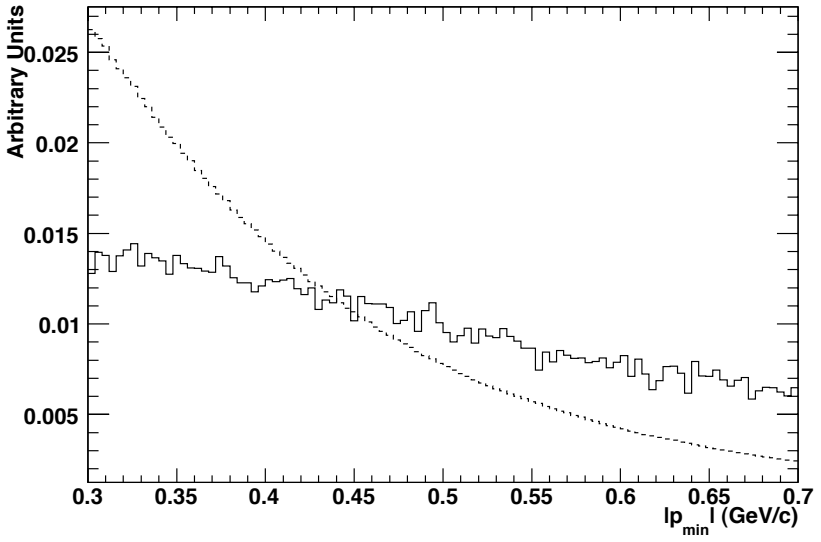


FIGURE 3.21: Normalised distributions of the momentum magnitude of the lowest momentum hadronic daughter for continuum (dashed line) and correctly reconstructed candidates (solid line) before post reconstruction cuts.

The cut on the laboratory momentum of daughters of the $X_{s/d}$ candidate, $|p_{min}|$, is tightened during cut optimisation with the requirement that it must be at least as tight as the value of 0.3 GeV/c implemented in the filter `BToXdGammaFilter`. Figure 3.21 shows distributions of this variable comparing correctly reconstructed candidates and combinatoric background for the lowest momentum daughter used to fit that candidate. Combinatoric backgrounds tend to favour lower momenta relative to correctly reconstructed candidates.

Finally for charged pions the tightness of the PID selector imposed on tracks used to reconstruct X_d candidates is varied to investigate which gives the best performance accepting signal and rejecting backgrounds.

3.6.4 *B* Meson Candidate Cuts

It is possible to tighten cuts applied to m'_{ES} and ΔE during event selection to reduce combinatoric backgrounds. This is providing that there is no significant signal loss which might bias the subsequent fit to these variables. Therefore the analysis requires $-0.3 \leq \Delta E \leq 0.2\,\text{GeV}$ and $m_{ES} \geq 5.22\,\text{GeV}/c^2$ the choice of which is based on the experience of previous *BABAR* radiative penguin analyses [1, 30, 59].

3.6.4.1 Hadronic Mass Bin Variations

As described above the choice of upper limit in the non-resonant hadronic mass bin is constrained by computational limitations in processing the pre-reconstruction event filters. The filter `BToXdGammaFilter` only reconstructs X_d candidates. Due to the difference in pion and kaon mass it follows that X_s candidates close to the upper limit can have significantly different efficiencies when passed through this filter. This is because the filter assumes pion mass for all charged tracks, whereas fully reconstructed X_s candidates redefine the mass of the track identified as a kaon and therefore have an increased reconstructed hadronic mass. In post-reconstruction event reduction the upper limit is therefore not allowed to exceed $2.0\,\text{GeV}/c^2$ thus minimising potential systematic differences between X_d and X_s candidate efficiencies due to this cut.

Due to increased combinatoric backgrounds with larger values of m_X the high mass region was previously observed to be statistically limited relative to the low mass resonant region despite it containing a larger fraction of the expected inclusive width [1]. Additionally modes where the hadronic component has a larger number of daughters or π^0 mesons will contribute more to combinatoric backgrounds. For this reason all cuts are optimised independently for six distinct classes in which both the upper mass limit and number of modes reconstructed in the high mass region are varied. Table 3.11 lists these classes and shows the total number events for the signal $B \to X_d\gamma$ MC data with with at least one reconstructed X_d candidate for each class before post reconstruction selections.

In the low mass resonant region there are three modes which dominate the signal contribution: 1,2 and 4 for ρ^0, ρ^+ and ω transitions respectively. To investigate the background contributions from including further reconstructed modes in this region the cuts

Class	Mass Limit (GeV/c^2)	Included Modes	MC X_d Evt ≥ 1 (True) X_d Cand
1	1.8	1,2,3,4	202,696 (75,347)
2	2.0	1,2,3,4	231,639 (84,816)
3	1.8	1,2,3,4,5,6,7	371,307 (98,582)
4	2.0	1,2,3,4,5,6,7	432,987 (112,752)
5	1.8	1,2,3,4,5,6,7,8,9	429,513 (110,592)
6	2.0	1,2,3,4,5,6,7,8,9	490,169 (126,743)

TABLE 3.11: Signal $B \to X_d\gamma$ MC data in the high mass region with at least one X_d candidate after applying cuts on the upper limit of reconstructed hadronic mass and number of reconstructed modes relevant to each optimisation class. Shown in parentheses are the number events with a correctly reconstructed candidate. The mode numbers correspond to those given in table 3.1

Class	Included Modes	MC Evt ≥ 1 (True) X_d Cand: $\rho^0;\rho^+;\omega$
7	1,2,4	199,690; 163,569; 74,625 (190,712; 144,934; 38,480)
8	1,2,3,4,5,6,7	202,941; 171,091; 108,232 (190,712; 144,934; 38,637)

TABLE 3.12: Signal $B \to X_d\gamma$ MC data in the low mass region with at least one X_d candidate after applying cuts on reconstructed hadronic mass and number of reconstructed modes relevant to each optimisation class. Shown in parentheses are the number of events with a correctly reconstructed candidate.

are optimised independently for two classes which vary the number of reconstructed modes. Table 3.12 lists these classes again showing the number of events with at least one reconstructed X_d candidate for signal MC data.

3.6.5 Multivariate Classifier for Continuum Event Reduction

The dominant contribution to combinatoric backgrounds is from continuum events. Ideally cuts on the topological and tagging variables discussed above can be used to reduce these backgrounds; however, the discrimination between signal and background for the majority of these variables is individually limited due to largely overlapping distributions. Therefore the use of a multivariate classifier is considered. Such classifiers use a number of user-defined input variables to create a discriminator which can recognise higher order patterns in data and use them to classify signal and background events. The classifier is trained on a subset of signal and background data where the data type is known beforehand. After training it can then be applied to further events to test them for consistency with the signal hypothesis. Such techniques are becoming increasingly common in high energy physics data analysis and in particular the *BABAR* collaboration have found them a particularly useful tool. In all tests neural networks (NN) are found to give the best performance for this analysis with MC data. Other multivariate classifiers investigated with these data include boosted and bagged decision trees; only

the NN is discussed here. The following gives a general overview of NN discriminators, before describing the optimisation process implemented to find the best performing NN.

3.6.5.1 Overview of Neural Networks

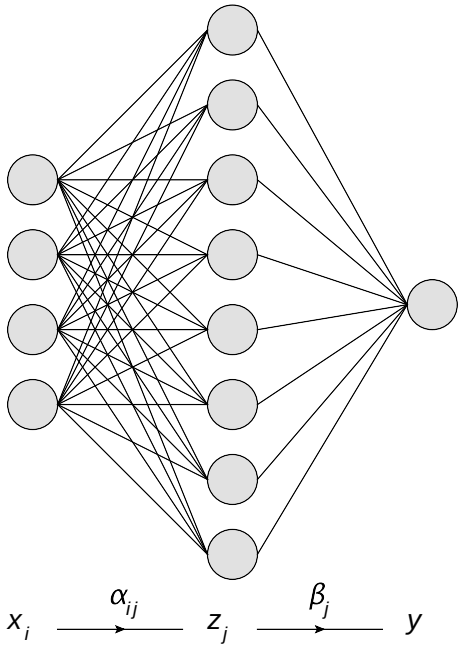


FIGURE 3.22: Schematic Representation of a Neural Network with four inputs, one hidden layer of eight nodes and one output.

Figure 3.22 shows a schematic representation of a typical NN with four input variables, one hidden layer with eight nodes and one output variable. The input variables, x_i, are normalised to the range $[0, 1]$ and are related to each node, j, in the hidden layer by a weight α_{ij} as represented by the lines in figure 3.22. The hidden layer then has a value z_j given by the output of an activation function, $\mathcal{A}$, whose input is the weighted sum of the input variables:

$$z_j = \mathcal{A}_j \left(\sum_{i=1}^{4} \alpha_{ij} x_i \right). \tag{3.12}$$

Similarly the output, y, is given by,

$$y = \mathcal{A}_y \left(\sum_{j=1}^{8} \beta_j z_j \right), \tag{3.13}$$

where β_j are weights between the hidden nodes and the output. Each training event has an associated quadratic classification error,

$$\epsilon = (Y - y)^2, \tag{3.14}$$

Variable	Cut Value
π^0 veto upper limit	$\geq 0.15\,\mathrm{GeV}/c^2$
π^0 veto lower limit	$\leq 0.1\,\mathrm{GeV}/c^2$
η veto upper limit	$\geq 0.59\,\mathrm{GeV}/c^2$
η veto lower limit	$\leq 0.5\,\mathrm{GeV}/c^2$
S'_{bump}	≤ 0.2
x_{trk}	$\geq 25\,\mathrm{cm}$
Lateral Moment	≤ 0.6
X_d $\lvert p_{min}\rvert$	$\geq 0.3\,\mathrm{GeV}/c$
X_d vertex χ^2 prob	≥ 0.01
Pion PID Selector	`pionKMTight`
Use Endcap Photons	No

TABLE 3.13: Unoptimised cut values used to assign training events.

where Y is the true event class, typically one for signal and zero for background. For each training event the network weights are systematically updated by propagating the previous quadratic error back from the output layer to the input layer. After all training events are exhausted one training cycle has been completed. An optimal NN will have n training cycles where n minimises the average quadratic error of a separate validation sample. This analysis uses the implementation of NN from the package `StatPatternRecognition` (SPR) [60], in which $\mathcal{A}$ is a sigmoid function.

3.6.5.2 Classification of Training and Testing Events

Training of the NN is carried out on continuum events which are as signal-like as possible. It is therefore desirable to apply all other post-reconstruction cuts beforehand. As the NN must be trained before cut optimisation a value for each cut to be optimised is chosen based on the experience of the previous *BABAR* analysis [1]. Table 3.13 lists the cut values chosen. After these cuts are applied the total efficiencies for simulated data are used to identify what proportion of these samples should be assigned to training to reflect the relative proportion of such events expected in data. Table 3.14 lists the required make up of training samples to reflect the expected proportion in data and indicates the proportion of events in each sample which must be assigned to training to obtain at least 10^4 training events for each of the optimisation classes.

Training events are assigned such that they are evenly sampled across the MC data to ensure they accurately reflect varying background conditions from different running periods which are emulated in the simulation. A second mutually exclusive dataset sampled in the same way is used for evaluating the performance of NNs as they are

Data Samples	Train Sample	Fraction of MC data
High Mass X_d Signal $X_{dd}; X_{du}$	1:1	0.25; 0.25
High Mass X_d Continuum uds; cc	5:1	0.043; 0.041
Low Mass X_d Signal $\rho^0; \rho^+; \omega$	4:3:1	0.053; 0.059; 0.111
Low Mass X_d Continuum uds; cc	4:1	0.111; 0.111

TABLE 3.14: Required make up of training samples to reflect expected proportions in data and the required fraction of MC data needed to give greater than 10^4 training events in each sample for each optimisation class.

Class	Training Events (Sig; Cont)	Testing Events (Sig; Cont)
1	11,479; 11,977	11,412; 11,904
2	12,984; 14,970	12,970; 14,898
3	13,664; 21,813	13,615; 21,722
4	15,624; 28,314	15,557; 28,205
5	15,233; 29,462	15,162; 29,363
6	17,485; 37,929	17,425; 37,832
7	15,476; 14,333	15,421; 14,244
8	15,478; 20,036	15,428; 19,928

TABLE 3.15: Size of NN training and testing samples for each optimisation class. There are two samples in each case, a sample of correctly reconstructed candidates and a sample of continuum backgrounds.

trained. For continuum events where there is more than one candidate in an event the candidate passing the requirements of best candidate selection described in section 3.6.6 is chosen.

For data analysis carried out after a NN cut has been applied, the use of events which trained the NN can lead to bias in efficiency calculations and fit studies. This is a consequence of potential over-training of the NN whereby the training dataset will give a better response than events not used for training. To eliminate this bias, training events are vetoed from the cut optimisation and subsequent studies; a correction factor is applied to account for this when calculating the efficiencies of signal event reconstruction. In signal MC data it is necessary to assign a consistent proportion of events which do not contain a correctly reconstructed candidate as training events even though they are not used in the actual training of the NN. This ensures that the correction applied to these data is the same for both correctly reconstructed candidates and combinatoric backgrounds; table 3.15 lists the size of test and training samples used for each optimisation class.

Variable Type	Considered Inputs
Topological	T_{ROE}, $\lvert\cos\theta_T\rvert$, R'_2, L_2^T, L_3^T, P_1^T, P_2^T, P_3^T, L_1^γ, L_2^γ, L_3^γ, P_1^γ, P_2^γ, P_3^γ, S_{ROE}, A_{ROE}, P_{ROE}, $\lvert\cos\theta_B\rvert$
Tagging	L_{TAG}^e, L_{TAG}^μ, L_{TAG}^l, H_{TAG}^π, H_{TAG}^K, $H_{TAG}^{K\pi}$, P_{TAG}

TABLE 3.16: Variables considered for NN input.

3.6.5.3 Training Strategy and Input Variable Selection

There are 25 topological and tagging variables considered for the NN classifier which are listed in table 3.16. Note that from the definitions given in sections 3.3.2.1 and 3.3.2.2, $T_{ROE} \equiv L_1^T$, thus only T_{ROE} is considered for NN input. Initially the structure of the NN is required to have one hidden layer with twice as many nodes as the input layer. For simplicity in implementing the analysis framework only one NN structure is optimised using the data corresponding to optimisation class 3. The choice of these data is arbitrary, input variable distributions are not observed to change significantly between classes and so the choice of data used to train and test is not expected to significantly affect NN optimisation. Once the optimal NN is found it is then trained using each of the datasets listed in table 3.15, only the number of training cycles varies between optimisation classes.

When training a NN the training data are randomly split into two sub samples of equal size. One sample is used to train the NN and the other to calculate the average quadratic error every ten training cycles. Once the number of training cycles with minimum quadratic error is determined the NN is retrained with that many cycles using all training data and its performance evaluated. The performance evaluation of a given NN uses SPR tools with the testing data sample to calculate the percentage of background events accepted for a cut on the NN output which accepts 40, 30, 20 and 10% of signal events. The calculated background acceptance for a given signal acceptance point has an associated statistical error. If the NN, when compared to an alternative NN, is found to have reduced background acceptance for one or more of the signal acceptance points the NN is considered to perform better. This is providing the improvement in background rejection percentage is greater than the calculated statistical error and that there is no significant degradation in background rejection at any of the other signal acceptance points.

SPR performs an internal calculation when training a NN, ranking the importance of the input variables. This calculation is not expected to be reliable but instead is used as a starting point when considering which variables to include. The NN is trained using all 26 input variables and the three highest ranked variables are then chosen to train the NN to be optimised. Variables are then added and removed in an arbitrary way to

evaluate which combination gives the best performance. If the addition or removal of a variable results in a better performing NN, determined by the criteria given above, then the resulting combination of input variables becomes the new baseline for adding and removing further variables. If no improvement in performance is observed then the pre-existing combination of variables forms the baseline. If performance does not deteriorate on removing a variable then preference is given to the NN with fewer input variables. The optimum is found to have ten input variables: T_{ROE}, R_2', L_3^T, L_2^γ, A_{ROE}, $|\cos\theta_B|$, L_{TAG}^e, L_{TAG}^μ, H_{TAG}^π and H_{TAG}^K. The number of hidden layers and corresponding nodes are again varied in an arbitrary way using these input variables. No NN is found to give better performance under these variations so the optimal NN is therefore determined to have one hidden layer with 20 nodes. The output of the optimal NN for both signal and continuum data not used for training is shown in figure 3.23 for event class three.

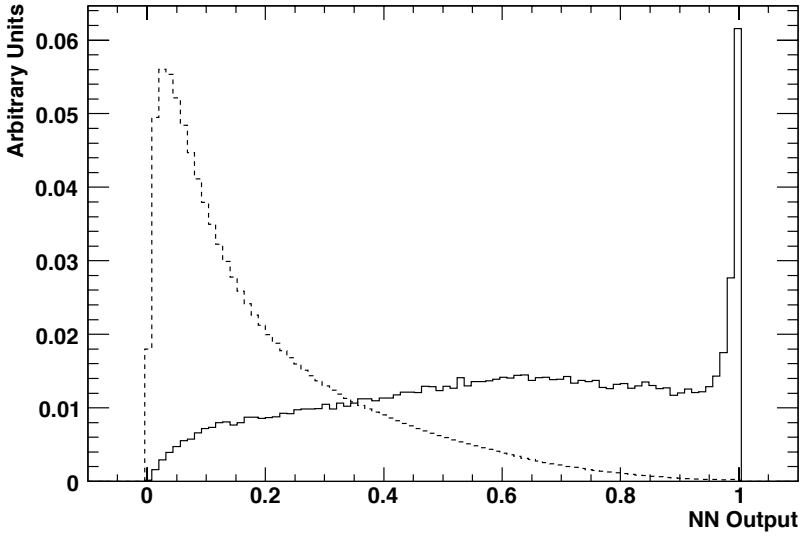


FIGURE 3.23: Normalised distributions of the NN output for continuum (dashed line) and correctly reconstructed candidates (solid line). These are MC data from event class three.

3.6.6 Best Candidate Selection

Once cuts have been applied there can still be events with more than one candidate. The likelihood fit described in chapter 4 assumes only one candidate per event so in the case of ambiguity a strategy for choosing which candidate to include in the fit is required. This selection must not bias the fit so the variables to be fitted, ΔE and m_{ES}', or any variable correlated with them cannot be used for candidate selection.

There are two steps to choosing the best candidate. Firstly the best candidate within a given reconstructed mode is chosen. Secondly if an event has more than one reconstructed mode then a preferred mode must be selected. For the former the following criteria are applied. If a mode has a single neutral particle (π^0 or η) then the candidate whose neutral daughter photon pair has invariant mass closest to the world average is chosen. For modes with two π^0 mesons this principle is extended into two dimensions by calculating the quantity,

$$\Delta m_{\pi^0} = \sqrt{(m^1_{\pi^0} - m^{WA}_{\pi^0})^2 + (m^2_{\pi^0} - m^{WA}_{\pi^0})^2}, \tag{3.15}$$

where $m^i_{\pi^0}$ is the invariant mass of the photon pair corresponding to the ith π^0 and $m^{WA}_{\pi^0}$ is the world average π^0 mass. The candidate with Δm_{π^0} closest to zero is chosen. If any ambiguity still exists or if a mode has no neutral daughters then the candidate with the highest vertex χ^2 probability is chosen. For modes where ambiguity exists and there is only one charged daughter then the candidate whose track has the largest χ^2 probability associated with its Kalman fit is chosen.

Investigations of signal MC data for patterns which may indicate an order of preference for choosing between candidates of different modes finds no preferred order so in the case of events where such ambiguity exists a random candidate is chosen.

3.6.7 Cut Optimisation

There are seven cuts for which simulated data are used to find an optimal value: the upper and lower limits of π^0 and η invariant mass vetoes; the upper limit on S'_{bump}; the lower limit on $|p_{min}|$; and the lower limit on the NN output. The choice of charged pion PID selector and use of endcap photons are also considered. Optimisation uses a simple cut and count technique in a signal region defined as m'_{ES} >5.27 GeV/c^2 and $-0.15 \leq \Delta E \leq 0.1$ GeV. The figure of merit (FOM) is $s/\sqrt{b}$ where s and b are the estimated luminosity normalised contributions in the signal region of correctly reconstructed events and combinatoric backgrounds respectively. The normalisation of MC data to the experimental luminosity is discussed in section 3.7.1. This FOM is preferred to $s/\sqrt{s+b}$ to ensure it does not directly depend on correctly normalising the signal $B \to X_d\gamma$ MC data; the contribution of incorrectly reconstructed $B \to X_d\gamma$ events to b is small.

The optimisation is iterative. From a baseline of cuts, an individual iteration makes systematic variations in each cut; all other cuts are held constant and the cut under consideration is varied. Continuous variables are each varied over a range of 20 values, the start value and step size of which are shown in table 3.17. Ranges are chosen to

Cut	Start Value	Step
upper π^0 veto	0.110	+0.005
lower π^0 veto	0.150	−0.005
upper η veto	0.530	+0.005
lower η veto	0.550	−0.005
S'_{bump}	0.3	−0.01
$\lvert p_{min} \rvert$	0.300	+0.025
NN out	0.60	+0.02

TABLE 3.17: Cut optimisation start point and step size for each continuous variable.

allow a significant proportion of phase space for each cut to be examined. For discrete variables the pion selector is varied to increasing levels of tightness and endcap photons are first considered and then vetoed. In each event, for each optimisation point, the best candidate selector is called if necessary then an event is tested for consistency with the signal region. The event is counted if it passes the cuts required for that optimisation point and is in the signal region. Correctly reconstructed candidate events contribute to s while all potential background candidates contribute to b. Thus after normalising each MC data type to reflect the expected contribution in experimental data, a given optimisation point has an associated FOM. The value giving the best FOM for the cut under consideration is used as the baseline for the next iteration. In the first iteration the baseline corresponds to the cut values in table 3.13 and requires a candidate to have associated NN output greater than 0.8.

The iterative process is continued until all cuts converge. If no convergence is observed after five iterations the set of cuts giving the best overall FOM is chosen. Table 3.18 lists the final FOM for each of the eight optimisation classes. The optimisation is performed independently for each event class.

This is a uni-directional optimisation process and is therefore only guaranteed to find a local maximum in the FOM; there may be an improved figure of merit found by changing the order in which variables are investigated. However, this analysis uses the cuts of the previous analysis [1] as a baseline which corresponds to a FOM of ~ 1.5 in the high mass region. These cuts were chosen using a similar FOM, optimised with an alternative method using independent MC data. Consequently as the optimisation here gives an improved result compared this baseline it is considered robust.

3.6.7.1 Discussion of Analysis Strategy

The cut optimisation results presented in table 3.18 show that in the high mass region class 2, with four reconstructed modes and an upper hadronic mass limit of $2.0\,\mathrm{GeV}/c^2$,

Class	s	b	$s/\sqrt{b}$
1	31.0	175.6	2.34
2	25.1	109.4	2.40
3	19.3	84.7	2.10
4	31.2	232.4	2.05
5	26.3	198.9	1.86
6	21.0	129.3	1.85
7	22.9	14.1	6.10
8	38.3	61.4	4.89

TABLE 3.18: Cut optimisation results for each optimisation class.

gives the most statistically significant measurement. In the low mass region the larger FOM is associated with class 7 which only reconstructs three modes. It follows that a measurement designed to maximise statistical significance should reconstruct fewer modes. However, following such a strategy leads to greater systematic uncertainty from model dependence when extrapolating to inclusive BFs. Clearly a balance must be struck between these competing factors and the decision of which optimisation class to use in each mass region must reflect this.

For the high mass region this analysis is required to be at least as inclusive as the previous *BABAR* measurement [1] which used optimisation class 3; this reconstructs seven modes with an upper hadronic mass limit of $1.8\,\mathrm{GeV}/c^2$. The degradation in $s/\sqrt{b}$ between classes 3 and 4 is $<\sim 3\%$ which is deemed to be an acceptable drop given the more inclusive nature of class 4 which reconstructs the same seven modes but over a larger hadronic mass range. However, moving to class 5 or 6 in lieu of class 3, i.e. including modes 8 and 9 which reconstruct two π^0 candidates, gives a decrease of $>\sim 10\%$ in the statistical FOM. It is also noted that despite their more inclusive nature these classes have smaller s due to cut optimisation points changing significantly[5] as a result of increased combinatorics from modes 8 and 9. In contrast class 4 has the largest value of s. Additionally modes 8 and 9 are susceptible to the $K^0_S \to \pi^0\pi^0$ background discussed in section 3.6.1. No specific reduction of this background is made and therefore, as such background is almost indistinguishable from signal, its presence can potentially bias the subsequent fit to data. The inclusion of modes 8 and 9 is therefore also undesirable for this reason. This analysis chooses optimisation class 4 in the high mass region. The upper hadronic mass limit is increased to $2.0\,\mathrm{GeV}/c^2$ compared to $1.8\,\mathrm{GeV}/c^2$ in the previous analysis; however, modes with two π^0 candidates in the hadronic reconstruction are not included in the measurement.

[5] Favouring tighter values.

Variable	Cut Value
Hadronic mass upper limit	$\leq$2.0 GeV/c^2
Included modes	1,2,3,4,5,6,7
m_{ES}	$\geq$5.22 GeV/c^2
ΔE	$\leq$0.2 GeV and ≥-0.3 GeV
π^0 veto	$\geq$0.155 GeV/c^2 and $\leq$0.095 GeV/c^2
η veto	$\geq$0.565 GeV/c^2 and $\leq$0.53 GeV/c^2
S'_{bump}	$\leq$0.2
x_{trk}	$\geq$25 cm
Lateral Moment	$\leq$0.6
X_d $\lvert p_{min}\rvert$	$\geq$0.425 GeV/c
X_d vertex χ^2 prob	$\geq$0.01
Pion PID Selector	`pionKMTight`
Use Endcap Photons	No
NN output	$\geq$0.86

TABLE 3.19: Optimised cut values in the high mass region for optimisation class 4.

The low mass region is dominated by resonant decays and the model used to simulate it assumes no other contributions. This accounts for the $\sim 20\%$ degradation in the FOM in moving from class 7 to class 8 which reconstructs four extra modes, none of which contribute to s in a purely resonant model. If the low mass region is reconstructed as entirely resonant contributions without allowing for the other modes reconstructed in the high mass region then the assumptions of quark hadron duality may break down. In particular the feed-through of high mass contributions across the threshold boundary due to detector resolution is not accounted for as required in the resonance modified KN model [27]. This results in the theoretical uncertainty when extracting $|V_{td}/V_{ts}|$ potentially being underestimated in chapter 1. Therefore, although the low mass region is dominated by resonant transitions, its measurement should be as inclusive as possible so as not to invalidate the subsequent extraction of CKM matrix parameters. For this reason the analysis chooses class 8 in this region; the same seven modes reconstructed in the high mass region are also all reconstructed in the low mass region thus allowing any potential non-resonant contributions to be measured. Tables 3.19 and 3.20 show the final cut values determined for optimisation classes 4 and 8 respectively.

3.7 Event Selection Summary

Following the choice of analysis strategy and determination of optimum cut values the subsequent event selection efficiencies and projected yields in experimental data are

Variable	Cut Value
Included modes	1,2,3,4,5,6,7
m_{ES}	$\geq$5.22 GeV/c^2
ΔE	$\leq$0.2 GeV and ≥ -0.3 GeV
π^0 veto	$\geq$0.16 GeV/c^2 and $\leq$0.11 GeV/c^2
η veto	$\geq$0.56 GeV/c^2 and $\leq$0.52 GeV/c^2
S'_{bump}	$\leq$0.19
x_{trk}	$\geq$25 cm
Lateral Moment	$\leq$0.6
X_d $\lvert p_{min}\rvert$	$\geq$0.6 GeV/c
X_d vertex χ^2 prob	$\geq$0.01
Pion PID Selector	`pionKMTight`
Use Endcap Photons	No
NN output	$\geq$0.92

TABLE 3.20: Optimised cut values in the low mass region for optimisation class 8.

presented. The final number of events with a reconstructed $B \to X_{s/d}\gamma$ candidate in MC data after vetoing both NN training events and any $B \to X_s\gamma$ events in generic B MC data are shown in table 3.21. The total number of $B \to X_d\gamma$ candidate events from $B \to X_s\gamma$ MC data with an intermediate $K^0_S \to \pi^+\pi^-$ decay which are correctly reconstructed are 1 and 6 in the high and low mass regions respectively. This shows that the selections applied in section 3.6.1 make such backgrounds negligible.

3.7.1 Estimated Data Yields and Selection Efficiencies

Table 3.21 can be used to estimate the expected yields for each sample in the experimental data. Estimated yields are calculated from the total number of $B\overline{B}$ pairs in experimental data (given in section 3.1.1) using the assumed BFs for signal modes and the calculated cross sections for $e^+e^- \to q\overline{q}$ ($q = u, d, s, c$) transitions at the $\Upsilon(4S)$ mass [4]. Correction factors to convert between MC data yields and estimated experimental data yields are calculated by taking the ratio of the expected data yield with the total MC dataset sizes listed in tables 3.3 and 3.4 for signal and table 3.2 for generic backgrounds. Signal MC data requires a further factor of two correction to account for the forced decay of one B meson in an event to a signal mode. In all calculations it is assumed that the transitions $\Upsilon(4S) \to B^+B^-$ and $\Upsilon(4S) \to B^0\overline{B}^0$ are equally likely. This assumption is consistent with current experimental measurements of $\Upsilon(4S) \to B\overline{B}$ transitions [14]; however, it means the uncertainty associated with such measurements is not considered a source of systematic uncertainty.

MC Data	High Mass X_d Events	Low Mass X_d Events	High Mass X_s Events	Low Mass X_s Events
KN $B \to X_d\gamma$	15,858 (7,972)	2,130 (93)	-	-
$B \to \rho^+\gamma$	7,481 (63)	24,551 (21,806)	-	-
$B \to \rho^0\gamma$	2,733 (29)	26,373 (26,027)	-	-
$B \to \omega\gamma$	3,784 (0)	4,667 (2,106)	-	-
KN $B \to X_s\gamma$	23,297 (945)	3,493 (316)	144,328 (87,978)	9,120 (474)
$B \to K^{*+}\gamma$	3,828 (0)	5,988 (992)	34,067 (212)	81,768 (73,626)
$B \to K^{*0}\gamma$	1,794 (0)	4,271 (2,315)	16,049 (306)	184,501 (182,665)
Generic B^+B^-	780	110	916	30
Generic $B^0\overline{B}^0$	532	45	664	39
Generic $c\overline{c}$	1,014	180	1,038	141
Generic $q\overline{q}$ $(q = u, d, s)$	1,763	419	882	108

TABLE 3.21: MC data sample sizes after event selection for both reconstructed $B \to X_d\gamma$ and $B \to X_s\gamma$ candidate events in the high and low mass region. Shown in parentheses are the number of correctly reconstructed candidates.

Sample	Fraction of σ_{BB} [4, 14]	Total Event Yield	Correction	High Mass Yield	MC Efficiency	Low Mass Yield	MC Efficiency
KN $B \to X_d\gamma$	8.98×10^{-6}	4,229	3.52×10^{-3}	74 (37)	0.88% (1.43%)	7 (0)	0.09% (0.01%)
$B^+ \to \rho^+\gamma$	3.85×10^{-7}	181	6.37×10^{-4}	5 (0)	1.32% (0.01%)	17 (15)	4.59% (4.08%)
$B^0 \to \rho^0\gamma$	4.05×10^{-7}	191	6.71×10^{-4}	2 (0)	0.48% (0.01%)	19 (19)	4.91% (4.85%)
$B^0 \to \omega\gamma$	2.30×10^{-7}	108	3.32×10^{-4}	1 (0)	0.58% (0.00%)	2 (1)	0.81% (0.40%)
KN $B \to X_s\gamma$	3.22×10^{-4}	151,724	0.0146	340 (14)	0.11% (0.02%)	51 (5)	0.02% (0.01%)
$B^+ \to K^{*+}\gamma$	1.89×10^{-5}	8,900	3.13×10^{-3}	12 (0)	0.07% (0.00%)	19 (3)	0.10% (0.05%)
$B^0 \to K^{*0}\gamma$	1.89×10^{-5}	8,900	3.11×10^{-3}	6 (0)	0.03% (0.00%)	13 (7)	0.07% (0.06%)
Generic B^+B^-	0.50	235.45×10^6	0.3322	259	(1.10×10^{-4})%	37	(1.55×10^{-5})%
Generic $B^0\overline{B}^0$	0.50	235.45×10^6	0.3279	174	(7.41×10^{-5})%	15	(6.27×10^{-6})%
Generic $c\bar{c}$	1.30	612.17×10^6	0.5424	574	(9.37×10^{-5})%	110	(1.79×10^{-5})%
Generic $q\overline{q}$ $(q = u, d, s)$	2.20	1035.98×10^6	0.6232	1,148	(1.11×10^{-4})%	294	(2.84×10^{-5})%

TABLE 3.22: Estimated X_d candidate yields and selection efficiencies. For each MC dataset the total number of events expected in data is given and the corresponding correction factor to convert between MC data and experimental data yields shown. The expected data yields after event selection are shown for both mass regions and the total efficiency given, where necessary these are corrected for NN training event vetoes. Parentheses indicate the expected yield of correctly reconstructed candidates.

Sample	High Mass Yield	MC Efficiency	Low Mass Yield	MC Efficiency
KN $B \to X_s\gamma$	2,107 (1,284)	0.69% (1.51%)	133 (7)	0.04% (0.01%)
$B^+ \to K^{*+}\gamma$	107 (1)	0.59% (0.01%)	256 (230)	1.42% (3.84%)
$B^0 \to K^{*0}\gamma$	50 (1)	0.28% (0.01%)	574 (568)	3.22% (4.79%)
Generic B^+B^-	304	(1.29×10^{-4})%	10	(4.23×10^{-6})%
Generic $B^0\overline{B}^0$	218	(9.25×10^{-5})%	13	(5.43×10^{-6})%
Generic $c\bar{c}$	563	(9.20×10^{-5})%	76	(1.25×10^{-5})%
Generic $q\bar{q}$ $(q = u, d, s)$	550	(5.31×10^{-5})%	67	(6.50×10^{-6})%

TABLE 3.23: Estimated X_s candidate yields and selection efficiencies. The expected data yields after event selection is shown for both mass regions and the total efficiency given. Parentheses indicate the expected yield of correctly reconstructed candidates.

Mode	X_d Events	Efficiency	X_s Events	Efficiency
1	2,570	3.45%	34,580	3.56%
2	2,204	2.65%	15,486	2.60%
3	1,185	1.65%	19,802	1.80%
4	1,198	0.81%	10,695	0.80%
5	156	0.47%	1,784	0.45%
6	398	0.29%	2,968	0.30%
7	261	2.72%	2,663	3.13%

TABLE 3.24: Generated events and total efficiency by mode for correctly reconstructed high mass X_d and X_s candidates.

Mass	X_d Events		Efficiency	X_s Events		Efficiency
(GeV/c^2)	Generated	Post-Cuts		Generated	Post-Cuts	
1.0-1.2	134,444	1,901	1.89%	802,481	19,511	2.43%
1.2-1.4	149,463	2,012	1.79%	1,051,158	23,404	2.23%
1.4-1.6	136,874	1,694	1.65%	1,052,711	19,690	1.87%
1.6-1.8	110,729	1,363	1.64%	885,260	14,956	1.69%
1.8-2.0	80,139	958	1.59%	652,094	10,072	1.54%
>2.0	131,611	44	0.04%	1,039,348	345	0.03%

TABLE 3.25: Generated signal modes and total efficiency as a function of generated hadronic mass for correctly reconstructed high mass X_d and X_s candidates.

Table 3.22 shows the estimated experimental data yields for all data samples before event selection. The correction factors are presented and post-selection yields for $B \to X_d\gamma$ candidate events are calculated applying the correction factors to the MC data sample sizes listed in 3.21. Where necessary further corrections have been made to account for vetoed NN training events. Also shown are the overall efficiency calculations of the event selection on each MC sample[6]. The corresponding post selection estimated yields and efficiencies for $B \to X_s\gamma$ candidate events are presented in table 3.23. A detailed breakdown of the number of generated events in signal MC data corresponding to a reconstructed mode are presented in appendix B. These generated numbers are used to calculate mode by mode efficiencies which are presented in table 3.24 for both flavours of data in the high mass bin.

Finally table 3.25 shows the selection efficiency for correctly reconstructed $B \to X_d\gamma$ and $B \to X_s\gamma$ candidate events as a function of generated hadronic mass in the high mass region. The denominator in this efficiency ratio corresponds to all generated signal

[6]The signal efficiencies here do not necessarily directly correspond to those given in section 4.6.4 as the respective denominators in efficiencies used to calculate measured BF values only consider events with generated hadronic mass within the reconstructed mass range

modes in the corresponding bin of generated hadronic mass before event selections are applied.

Chapter 4

Signal Yield Extraction

Yields for the number of correctly reconstructed candidates in experimental data, after event selection, are extracted using a two dimensional (2D) unbinned extended maximum likelihood fit[1] to the kinematic variables ΔE and m'_{ES}. The fit strategy is optimised using simulated data samples. This chapter introduces the maximum likelihood fit (MLF) technique, before discussing the probability density functions (PDFs) used to model different event hypotheses and their parameterisation using MC data. The combined fitting strategy to be used on experimental data is outlined and studies for bias in this strategy using emulated and embedded MC data samples are presented. Finally the results of each of the four fits to the experimental data are given.

4.1 Maximum Likelihood

The MLF is a powerful statistical tool described in detail elsewhere [63, 64], an overview of which is given here. It can be advantageous compared with other fitting techniques, such as least squares, for it treats data on an event by event basis thus avoiding bias which can arise from binning samples beforehand.

4.1.1 Overview

Consider a sample of N events described by variable x whose parent distribution is a normalised PDF, $\mathcal{P}(x;\alpha_1,..,\alpha_n)$, described by n parameters, α_k $(k = 1,..,n)$. The sample can provide estimators for each parameter, $\hat{\alpha}_k$, through the principle of maximum likelihood (PML). For a given set of estimators the likelihood, $\mathcal{L}$, for the sample is given

[1] All computational fits use the `RooFit` [61] toolkit which implements the `Minuit` core package [62].

by,

$$\mathcal{L}(\hat{\alpha}_1, .., \hat{\alpha}_n) = \prod_{i=1}^{N} \mathcal{P}(x_i; \hat{\alpha}_1, .., \hat{\alpha}_n). \tag{4.1}$$

This represents the joint probability density of obtaining the x_i values observed[2]. The PML states that the set of $\hat{\alpha}_k$ which maximise $\mathcal{L}$ will best describe the data as these give a higher cumulative probability. Therefore iterative variation of estimators to find the maximum value of $\mathcal{L}$ provides a robust method of estimating the parameterisation of $\mathcal{P}(x; \alpha_1, .., \alpha_n)$, thus fitting the PDF to the data sample.

It is possible to generalise (4.1) such that a sample described by two independent variables, (x, y), and constructed from M event hypotheses, e.g. signal and different backgrounds, will have a likelihood function given by,

$$\mathcal{L}(\hat{\alpha}^x_1, .., \hat{\alpha}^x_n, \hat{\alpha}^y_1, .., \hat{\alpha}^y_m) = \prod_{i=1}^{N} \left[\sum_{j=1}^{M} w_j \mathcal{P}^x_j(x_i; \hat{\alpha}^x_1, .., \hat{\alpha}^x_n) \mathcal{P}^y_j(y_i; \hat{\alpha}^y_1, .., \hat{\alpha}^y_m) \right], \tag{4.2}$$

where w_j is the assigned weight of each hypothesis.

For computational calculations it is more convenient to find the best set of estimators by minimising the negative log-likelihood, l, as the product over N events in (4.1) becomes a sum,

$$l = -\ln \mathcal{L} = -\sum_{i=1}^{N} \ln \mathcal{P}(x_i; \hat{\alpha}_1, .., \hat{\alpha}_n). \tag{4.3}$$

The best values for each estimator are obtained from the set of simultaneous equations,

$$\frac{\partial l}{\partial \hat{\alpha}_k} = 0. \tag{4.4}$$

Corresponding uncertainties on estimator values are calculated by taking second derivatives of l, assuming the minimum is locally parabolic. The associated error matrix, E, is given by,

$$\left(E^{-1}\right)_{ij} = \left(\frac{\partial^2 l}{\partial \hat{\alpha}_i \partial \hat{\alpha}_j} \right). \tag{4.5}$$

4.1.2 Extended Maximum Likelihood

The above discussion concerns estimators calculated to determine the shape of data distributions where the PDFs are normalised such that,

$$\int \mathcal{P}(x; \alpha_1, .., \alpha_n)\, dx = 1. \tag{4.6}$$

[2]Omitted from the right hand side of (4.1) and subsequent equations are factors of $1/N!$ as ultimately it is variations in $\mathcal{L}$ which are of interest.

However, if the normalisation is also an unknown to be determined from the fit, the extended maximum likelihood fit (EMLF) is required. This method relaxes the constraint of (4.6) replacing $\mathcal{P}(x;\alpha_1,..,\alpha_n)$ with a function $\mathcal{Q}(x;\alpha_1,..,\alpha_n)$ such that,

$$\int \mathcal{Q}(x;\alpha_1,..,\alpha_n)\,dx = \nu, \tag{4.7}$$

where ν is the mean sample size. In finite data where events are sampled at random the actual sample size N will not necessarily coincide with ν. This is reflected in the EMLF by multiplying the likelihood by the Poisson probability of obtaining N events given a mean ν such that,

$$\begin{aligned} l &= -\ln\left(e^{-\nu}\nu^N\right) - \sum_{i=1}^{N} \ln \mathcal{P}(x;\alpha_1,..,\alpha_n), \\ &= \nu - \sum_{i=1}^{N} \ln \mathcal{Q}(x;\alpha_1,..,\alpha_n). \end{aligned} \tag{4.8}$$

As the normalisation of $\mathcal{Q}$ increases, the sum on the right hand side of (4.8) will become more negative; however, the opposite is true of the ν term. Thus, by balancing these effects, minimising l will find the most probable value of ν given the PDF parameterisation of x for a sample of N events.

4.1.3 Fit Variables

The choice of m'_{ES} and ΔE as fit variables aims to maximise statistical significance when fitting to experimental data. The use of two variables improves the ability of the fit to discriminate between different event hypotheses. Where possible event hypotheses are modelled as the product of uncorrelated PDFs following the formalism of (4.2). Consequently the variables used in the fit should be essentially independent. The variables m_{ES} and ΔE are to a good approximation independent for signal; however, as a consequence of (3.3) this is not necessarily true for m'_{ES} and ΔE. Despite this, previous *BABAR* analyses in the radiative penguin decay working group have shown that m'_{ES} and ΔE are suitable variables for a two dimensional MLF [1, 30, 59]. The improved resolution obtained from m'_{ES} and the need to be consistent with previous *BABAR* radiative penguin analyses therefore makes it the preferential variable to fit for over m_{ES}.

4.2 Corrections to $B \to X\gamma$ MC data

Previous measurements with *BABAR* experimental data of $B \to X_s\gamma$ transitions show that JETSET phase space decays do not accurately model the relative contributions

Fit	Yield	Pre-weighting	Post-weighting
$B \to X_s^L \gamma$	Signal	805	801
	Mass Cross-feed	126	162
$B \to X_s^H \gamma$	Signal	1284	1112
	Self Cross-feed	823	879
$B \to X_d^L \gamma$	PID Cross-feed	15	13
	X_s Mass Cross-feed	46	36
$B \to X_d^H \gamma$	PID Cross-feed	14	11
	X_s Self Cross-feed	326	324

TABLE 4.1: Corrected event yield estimates for $B \to X_s\gamma$ event hypotheses after MC data weighting.

from different final states [1, 30]. These analyses calculated weights (assumed to be independent of hadronic mass) which can be applied to different classes of MC events to make the model more representative of the experimental data. Corrections are not applied to resonant signal MC as K^*, ρ and ω decays are well measured [14].

The analysis in [30] reconstructed 38 exclusive $B \to X_s\gamma$ final states in the hadronic mass range $1.1 \leq m_X \leq 2.8\,\mathrm{GeV}/c^2$ using experimental data up to and including Run 2. For the purposes of weighting MC data these final states are classified into ten distinct sets. The analysis in [1] is the previous version of the analysis presented here and reconstructed modes 1-7 of table 3.1 in the hadronic mass range $1.0 \leq m_X \leq 1.8\,\mathrm{GeV}/c^2$ using experimental data up to and including Run 5. Summaries of final state contributions in generated $B \to X\gamma$ MC data are given in appendix B and the application of weights to $B \to X_s\gamma$ MC events is summarised in table B.6. Weights are applied such that the overall normalisation of MC data before cuts is constant resulting in a further correction weight to any mode not reconstructed in either of the above analyses to preserve unity. After weights have been applied the change in the relative contribution of individual events means the post event selection $B \to X_s\gamma$ MC data yields and thus expected experimental data yields must be corrected. Table 4.1 shows the corrected estimated experimental data yields. These are compared to the unweighted estimates given in tables 3.22 and 3.23. All subsequent plots and calculations pertaining to $B \to X_s\gamma$ MC events in the high mass region use the weighted data.

No correction is made to $B \to X_d\gamma$ data generated through the JETSET phase space model as information from experimental data does not exist; a naive application of the $B \to X_s\gamma$ weights to $B \to X_d\gamma$ MC events is performed in section 5.1.4.2 as part of the study of systematic uncertainties in this analysis.

4.3 Event Hypotheses

In fits to experimental data all potential event hypotheses must be identified and their relative contribution modelled. Here contributing events are determined from MC data. The respective high and low mass regions[3] for a given flavour of $B \to X\gamma$ candidates are modelled with identical event hypotheses. The following section presents all of the event hypotheses identified for this analysis.

4.3.1 $B \to X_s\gamma$ Candidate Events

Five contributing event hypotheses are identified for $B \to X_s\gamma$ candidates passing all event selection cuts. Signal MC data can contribute either correctly or incorrectly reconstructed events. Incorrectly reconstructed signal events are divided into those where the generated decay corresponds to the reconstructed mass region[4] (self cross-feed) and those where the generated decay was from the other mass region (mass cross-feed). The remaining two event hypotheses are any other $B\overline{B}$ event (generic $B\overline{B}$) and any continuum event. Figure 4.1 shows 2D plots of m'_{ES} and ΔE from MC data in the low mass region. The corresponding plots for the high mass region are shown in figure 4.2. Signal, generic $B\overline{B}$ and continuum event distributions can be sufficiently fitted as the product of independent PDFs for each variable. The same is not true for cross-feed events where the combination of different background contributions leads to correlations between the fit variables. For example signal decays with a π^0 in the final state have enhanced backgrounds in the signal region due to background candidates where a relatively low energy photon from the π^0 decay is exchanged for a similar background photon. The resulting candidate will be kinematically very similar to the generated signal candidate; this is particularly evident in low mass $B^+ \to K^{*+}\gamma$ self cross-feed events where the hadronic state is reconstructed from $K^{*+} \to K^+\pi^0$ decays. Such background does not exist in decays without a neutral hadronic final state daughter. Other contributions arise from backgrounds where the reconstructed candidate was not the same as the generated signal candidate or where the generated event was not a signal decay. In theory such correlations in these data can be more accurately modelled by identifying each background category and assigning it a separate event hypothesis; however, this is impractical as there are many such categories and the resulting combined fit to experimental data would become unmanageable. Consequently cross-feed distributions must be modelled as 2D histogram PDFs whose shapes are determined from MC data.

[3] To simplify notation, X^L will be used to indicate a hadronic candidate with reconstructed mass in the range $0.5 \leq m_X < 1.0\,\mathrm{GeV}/c^2$ and X^H a hadronic candidate with reconstructed mass in the range $1.0 \leq m_X \leq 2.0\,\mathrm{GeV}/c^2$

[4] Any cross-feed component from the high mass bin includes events where the generated hadronic mass was greater than $2.0\,\mathrm{GeV}/c^2$

The choice of binning for these histograms is arbitrary and depends on the quantity of MC data available; more data allow finer binning. In this analysis the statistics in all cross-feed MC data are sufficient to allow ten bins in each of the signal regions defined in the previous chapter for cut optimisation. This results in 20 bins for ΔE and 35 bins for m'_{ES} across the full fit range.

4.3.2 $B \to X_d\gamma$ Candidate Events

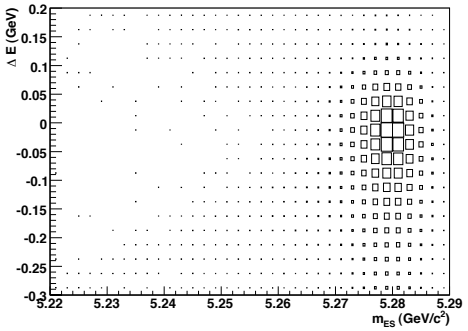
Eight event hypotheses are identified for $B \to X_d\gamma$ candidates passing all event selection cuts. Five of these directly correspond to the event hypotheses in $B \to X_s\gamma$ data: Correctly reconstructed events, self cross-feed events, mass cross-feed events, generic $B\overline{B}$ events and continuum events. The additional three event hypotheses correspond to cross-feed contributions where $B \to X_d\gamma$ candidates were reconstructed in $B \to X_s\gamma$ data. The three categories are correctly reconstructed $B \to X_s\gamma$ events where the generated charged kaon was reconstructed as a charged pion due to PID inefficiency (PID cross-feed) and incorrectly reconstructed events (separated into X_s self cross-feed and X_s mass cross-feed analogous to the signal cross-feed contributions). Figure 4.3 shows 2D plots of m'_{ES} and ΔE from MC data in the low mass region. The corresponding plots for the high mass region are shown in figure 4.4. Signal, PID cross-feed, generic $B\overline{B}$ and continuum event distributions can be sufficiently fitted as the product of independent PDFs in each variable. The remaining cross-feed events are again correlated due to combinations of different background contributions and must therefore be modelled as 2D histogram PDFs.

4.4 PDF Parameterisation

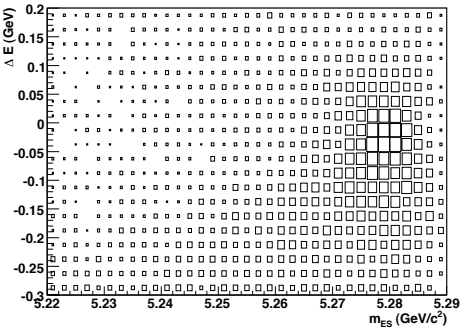
The choice of PDF parameterisation for different event hypotheses is determined from MLFs to MC data passing the event selection criteria. A set of PDFs is required for each of the four fits to be carried out. Presented here are the sets of PDFs which were found to best fit the MC data and their fitted parameters from those data.

4.4.1 PDF Functional Forms

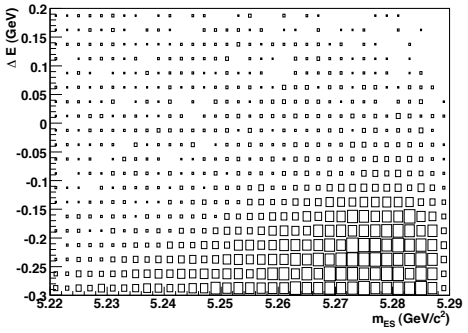
The functional form for all PDF distributions used in the subsequent MLFs are given here for a general variable x and normalisation factor $\mathcal{N}$. Each parameter to be fitted is introduced.



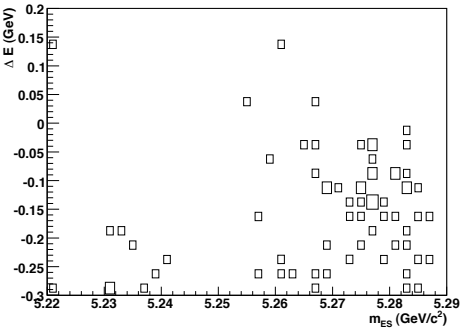
(a) $B \to K^*\gamma$ signal events with a correctly reconstructed candidate (signal).



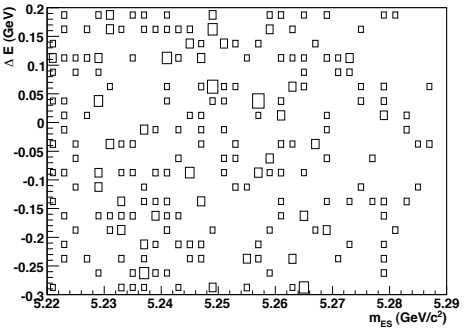
(b) $B \to K^*\gamma$ events where the candidate was incorrectly reconstructed (self cross-feed).



(c) $B \to X_s\gamma$ KN model events where the candidate was incorrectly reconstructed (mass cross-feed).

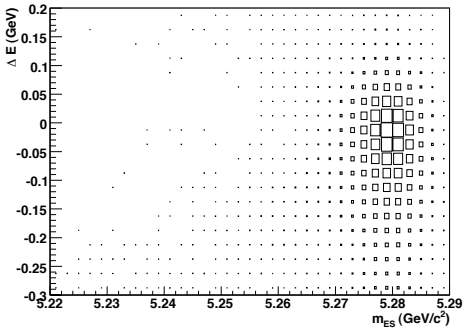


(d) Generic $B\overline{B}$ events.

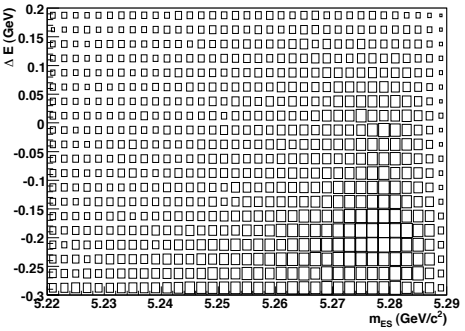


(e) Continuum events.

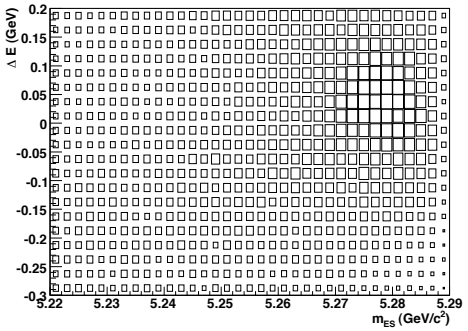
FIGURE 4.1: 2D plots of m'_{ES} and ΔE showing $B \to X_s^L\gamma$ candidates in each event hypothesis for MC data passing event selection cuts. The number of events in each plot reflects the MC data statistics given in table 3.21. Individual components in the generic $B\overline{B}$ and continuum data are weighted such that the combined data plots reflect any difference in relative contributions expected from those components in experimental data.



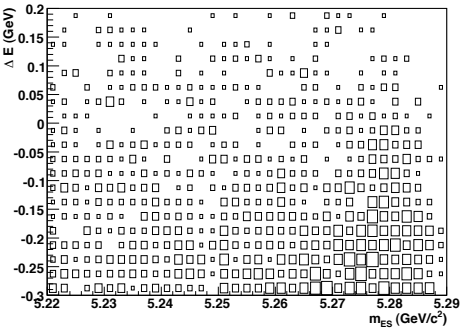
(a) $B \to X_s\gamma$ KN model events where the candidate was correctly reconstructed (signal).



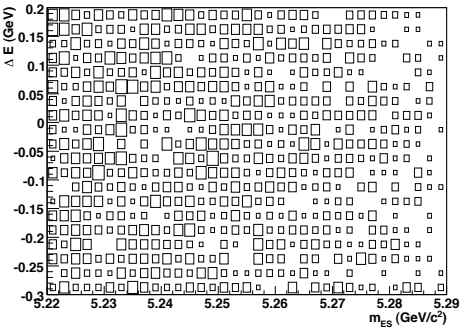
(b) $B \to X_s\gamma$ KN model events where the candidate was incorrectly reconstructed (self cross-feed).



(c) $B \to K^*\gamma$ events where the candidate was incorrectly reconstructed (mass cross-feed).

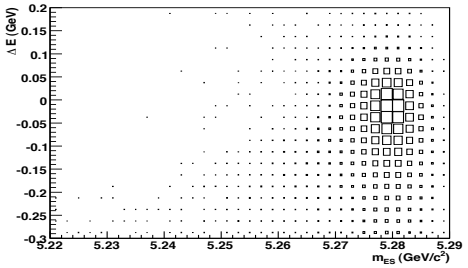


(d) Generic $B\overline{B}$ events.

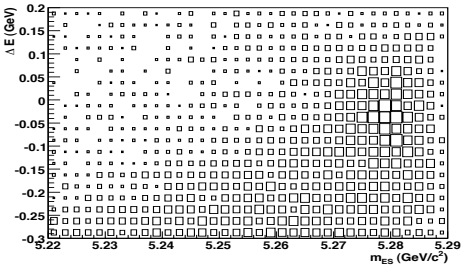


(e) Continuum events.

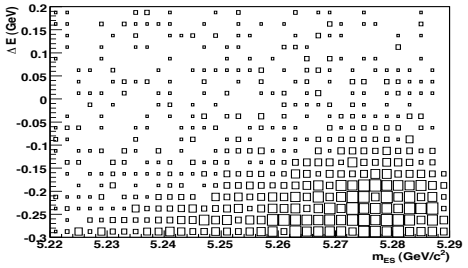
FIGURE 4.2: 2D plots of m'_{ES} and ΔE showing $B \to X_s^H\gamma$ candidates in each event hypothesis for MC data passing event selection cuts. The number of events in each plot reflects the MC data statistics given in table 3.21. Individual components in the generic $B\overline{B}$ and continuum data are weighted such that the combined data plots reflect any difference in relative contributions expected from those components in experimental data.



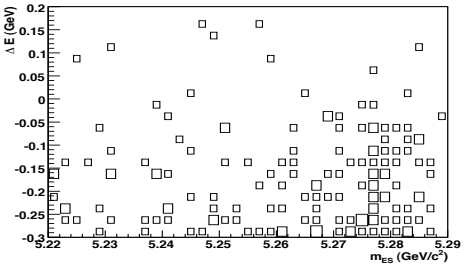
(a) $B \to \rho/\omega\gamma$ events where the candidate was correctly reconstructed (signal).



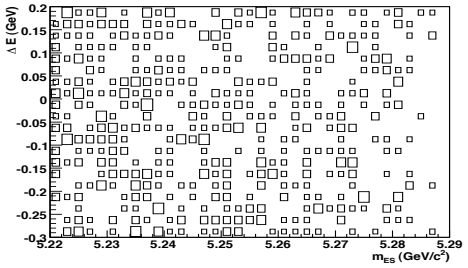
(b) $B \to \rho/\omega\gamma$ events where the candidate was incorrectly reconstructed (self cross-feed).



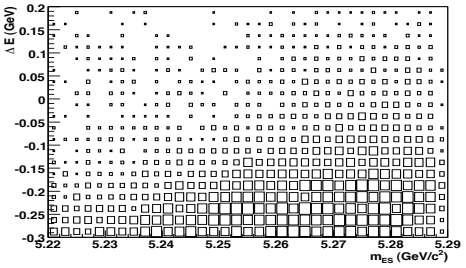
(c) $B \to X_d\gamma$ events where the candidate was incorrectly reconstructed (mass cross-feed).



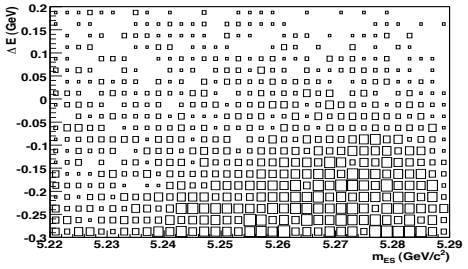
(d) Generic $B\overline{B}$ events.



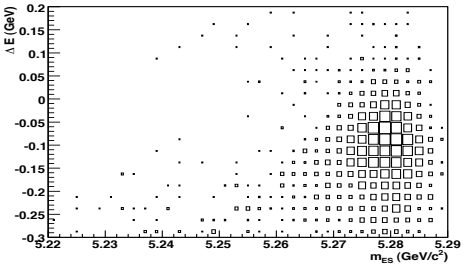
(e) Continuum events.



(f) $B \to K^*\gamma\ X_s$ events where the candidate was incorrectly reconstructed (X_s self cross-feed).

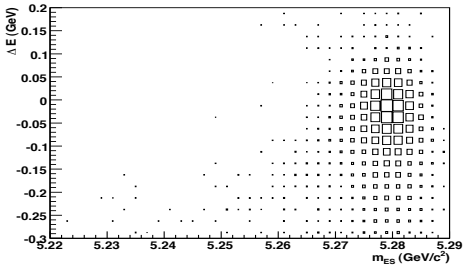


(g) $B \to X_s\gamma\ X_s$ KN model events where the candidate was incorrectly reconstructed (X_s mass cross-feed).

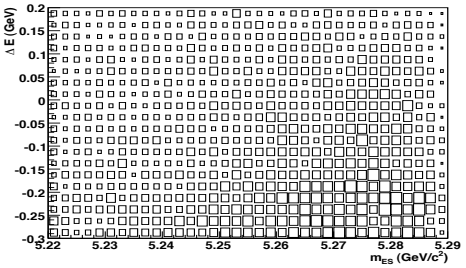


(h) $B \to X^s\gamma$ events where the candidate was correctly reconstructed but the charged kaon was incorrectly identified as a charged pion (X_s PID cross-feed).

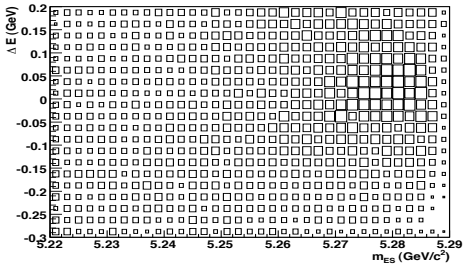
FIGURE 4.3: 2D plots of m'_{ES} and ΔE showing $B \to X_d^L\gamma$ candidates in each event hypothesis for MC data passing event selection cuts. The number of events in each plots reflects the MC data statistics given in table 3.21. Individual components in the generic $B\overline{B}$ and continuum data are weighted such that the combined data plots reflect any difference in relative contributions expected from those components in experimental data.



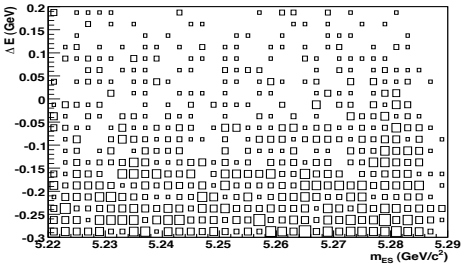
(a) $B \to X_d\gamma$ KN model events where the candidate was correctly reconstructed.



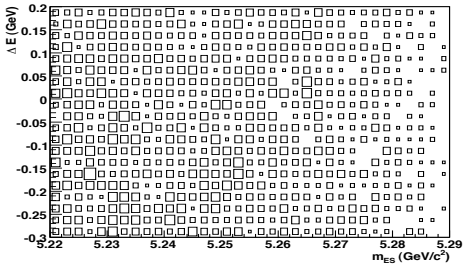
(b) $B \to X_d\gamma$ KN model events where the candidate was incorrectly reconstructed (self cross-feed).



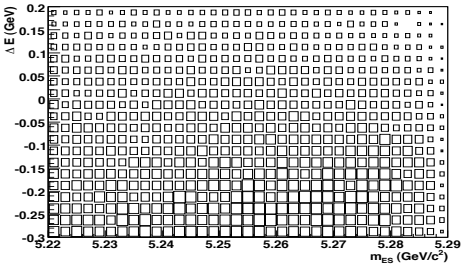
(c) $B \to \rho/\omega\gamma$ events where the candidate was incorrectly reconstructed (mass cross-feed).



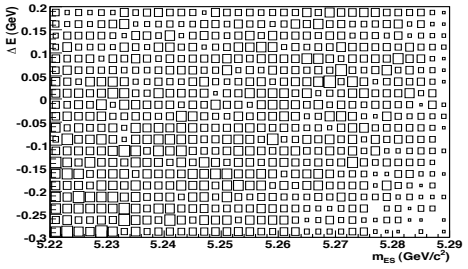
(d) Generic $B\overline{B}$ events.



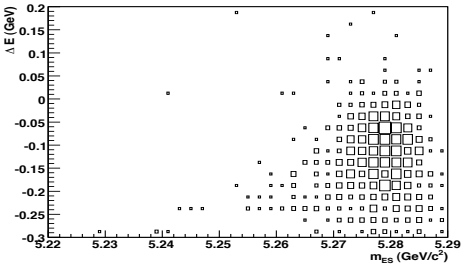
(e) Continuum events.



(f) $B \to X_s\gamma$ KN model events where the candidate was incorrectly reconstructed (X_s self cross-feed).



(g) $B \to K^*\gamma$ events where the candidate was incorrectly reconstructed (X_s mass cross-feed).



(h) $B \to X^s\gamma$ events where the candidate was correctly reconstructed but the charged kaon was incorrectly identified as a charged pion (X_s PID cross-feed).

FIGURE 4.4: 2D plots of m'_{ES} and ΔE showing $B \to X_d^H\gamma$ candidates in each event hypothesis for MC data passing event selection cuts. The number of events in each plots reflects the MC data statistics given in table 3.21. Individual components in the generic $B\overline{B}$ and continuum data are weighted such that the combined data plots reflect any difference in relative contributions expected from those components in experimental data.

The Argus function [65] is commonly used in high energy physics to model combinatoric backgrounds where the variable phase space has a kinematic threshold. It has the functional form,

$$\mathcal{P}(x; X_0, p, \xi) = \frac{1}{\mathcal{N}} x \left[F(x)\right]^p \exp\left[\xi F(x)\right], \quad F(x) = 1 - \left(\frac{x}{X_0}\right)^2, \tag{4.9}$$

where p is a power variable, ξ determines the distribution slope and X_0 represents an upper kinematic limit. In this analysis the Argus function models combinatoric backgrounds in m'_{ES}. The threshold is thus defined by the kinematic constraint from the beam energy which fixes X_0=5.29 GeV/c^2.

The Gaussian function models symmetric peaking distributions and has functional form,

$$\mathcal{P}(x; \mu, \sigma) = \frac{1}{\mathcal{N}} \exp\left(-0.5 \left(\frac{x-\mu}{\sigma}\right)^2\right), \tag{4.10}$$

where μ is the mean and σ the width of the distribution.

The Crystal-Ball (CB) [66] function modifies the Gaussian distribution by enhancing the low end tail and therefore improves the modelling of distributions where such a tail exists due to energy loss. It has the following form,

$$\mathcal{P}(x; \mu, \sigma, \alpha, n) = \frac{1}{\mathcal{N}} \begin{cases} \exp\left(-0.5 \left(\frac{x-\mu}{\sigma}\right)^2\right), & \text{for } \frac{x-\mu}{\sigma} > -\alpha \\ \left(\frac{n}{|\alpha|} - |\alpha| - \frac{x-\mu}{\sigma}\right)^{-n} \mathcal{A}(\alpha, n), & \text{for } \frac{x-\mu}{\sigma} \leq -\alpha \end{cases} \tag{4.11}$$

where $\mathcal{A}(\alpha, n) = (n/|\alpha|)^n \exp(-|\alpha|^2/2)$. The parameter α determines where the tail intersects the Gaussian distribution and the parameter n describes the shape of the tail.

The 'Cruijff' function [67] modifies the Gaussian distribution by allowing independent widths, σ_L and σ_R, either side of the mean and allows non-Gaussian tails either side of the mean through the parameters α_L and α_R. It has functional form,

$$\mathcal{P}(x; \mu, \sigma_R, \alpha_R, \sigma_L, \alpha_L) = \frac{1}{\mathcal{N}} \begin{cases} \exp\left(-\frac{(x-\mu)^2}{2\sigma_R^2 + \alpha_R (x-\mu)^2}\right), & \text{for } x \geq \mu \\ \exp\left(-\frac{(x-\mu)^2}{2\sigma_L^2 + \alpha_L (x-\mu)^2}\right), & \text{for } x < \mu \end{cases} \tag{4.12}$$

Finally the polynomial function has the form,

$$\mathcal{P}(x; a_1, .., a_N) = \frac{1}{\mathcal{N}} \left(1 + \sum_{n=1}^{N} a_n x^n\right). \tag{4.13}$$

The polynomial order, N, is chosen before the MLF.

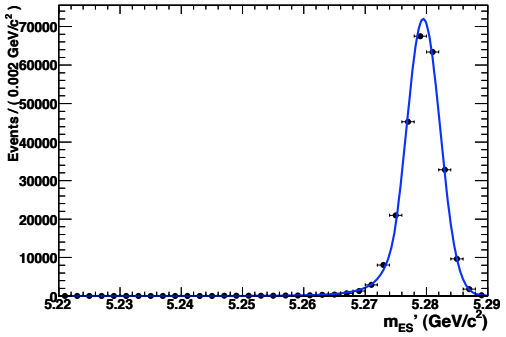
Hypothesis	Parameter	Low Mass Fit	High Mass Fit
Signal	μ (GeV/c^2)	5.27949±0.00001	5.27942±0.00001
	σ (GeV/c^2)	0.002741±0.000006	0.002755±0.000015
	α	1.613±0.013	1.647±0.026
	n	3.926±0.083	3.74±0.15
Generic $B\overline{B}$	f	0.42±0.17	0.19±0.11
	μ (GeV/c^2)	5.2793±0.0017	5.2793±0.0010
	σ (GeV/c^2)	0.0043±0.0017	0.0047±0.0013
	ξ	−117±82	−69±43
	p	0.67±0.80	0.67±0.71
Continuum	ξ	−54±29	−9.3±9.3
	p	1.04±0.33	0.553±0.098

TABLE 4.2: Fitted m'_{ES} PDF parameters in MC data for each event hypothesis showing results for both low and high mass region $B \to X_s\gamma$ candidates.

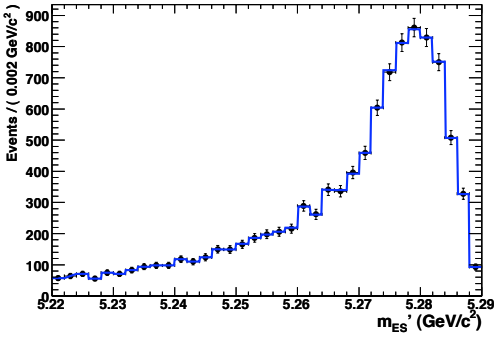
4.4.2 $B \to X_s\gamma$ Candidate PDFs

Signal $B \to X_s\gamma$ events are modelled as a CB function in m'_{ES} and a Cruijff function in ΔE as these PDFs allow for the asymmetries which arise in the signal distributions due to energy leakage. Self cross-feed and mass cross-feed events are modelled as 2D histogram PDFs whose shapes are determined from MC data. Generic $B\overline{B}$ events are modelled as the sum of a Gaussian and Argus function in m'_{ES}, with the Gaussian having relative normalisation, f, and a order-2 polynomial in ΔE. The Gaussian function component in m'_{ES} allows for an observed peak in the signal region which can include $B \to X_s\pi^0$ decays where the π^0 decayed asymmetrically and the soft photon was not reconstructed resulting in a failure of the π^0 veto used in event selection. Finally continuum events are modelled as an Argus function in m'_{ES} and a order-1 polynomial in ΔE.

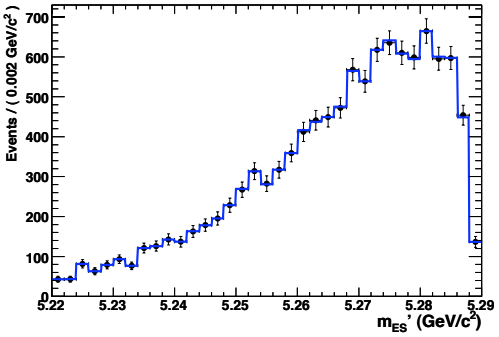
Figures 4.5 and 4.6 show the fits to MC data of m'_{ES} and ΔE respectively for $B \to X_s\gamma$ candidates in the low mass region and tables 4.2 and 4.3 list the respective estimator values and associated errors from each PDF fit. The corresponding fits for $B \to X_s\gamma$ candidates in the high mass region are shown in figures 4.7 and 4.8 for m'_{ES} and ΔE respectively with the fitted estimator values and their errors also listed in tables 4.2 and 4.3. Models constructed from 2D histogram PDFs are not fitted to the MC data, instead the MC data histogram fully determines the PDF. It follows that in plots showing these models the PDF and MC data points are entirely coincident.



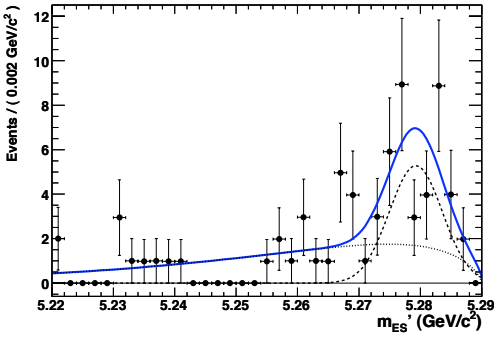
(a) $B \to K^*\gamma$ events with a correctly reconstructed candidate (signal) modelled as a CB function.



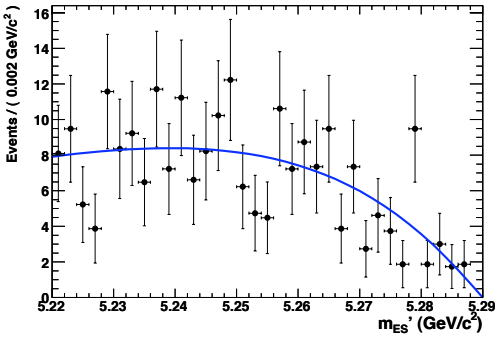
(b) $B \to K^*\gamma$ events where the candidate was incorrectly reconstructed (self cross-feed). The PDF is the 1D projection of the 2D histogram.



(c) $B \to X_s\gamma$ KN model events where the candidate was incorrectly reconstructed (mass cross-feed). The PDF is the 1D projection of the 2D histogram.

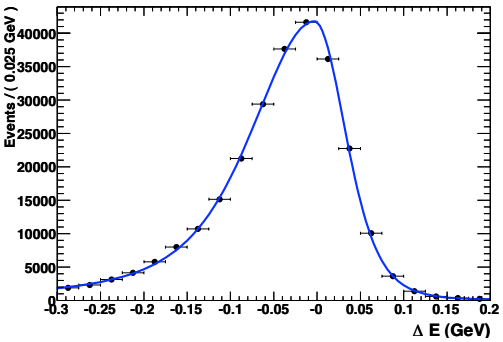


(d) Generic $B\overline{B}$ events modelled as the sum of a Gaussian (dashed) and Argus (dotted) function.

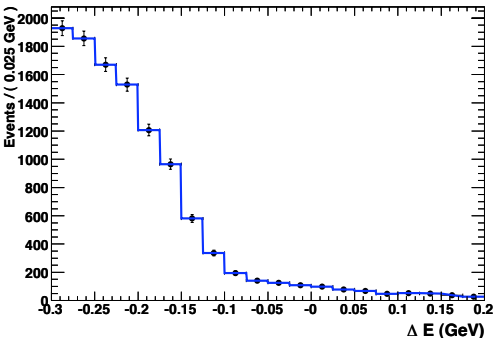


(e) Continuum events modelled as an Argus function.

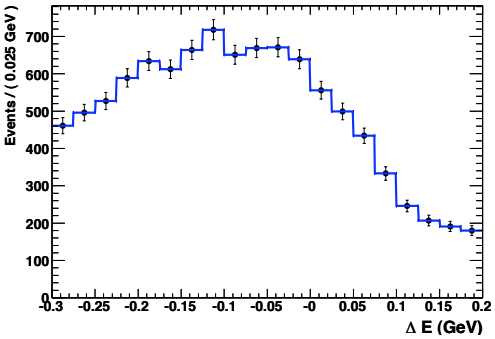
FIGURE 4.5: m'_{ES} distribution PDF fits to MC data for $B \to X_s^L\gamma$ candidates. PDF distributions (solid blue line) are overlaid on the MC data whose error bars reflect the statistical uncertainty from the chosen binning of those data. In all cases the number of events indicated by the ordinate is arbitrarily determined from the MC data statistics.



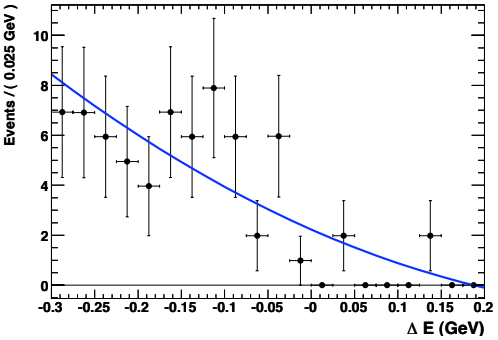
(a) $B \to K^*\gamma$ events with a correctly reconstructed candidate (signal) modelled as a Cruijff function.



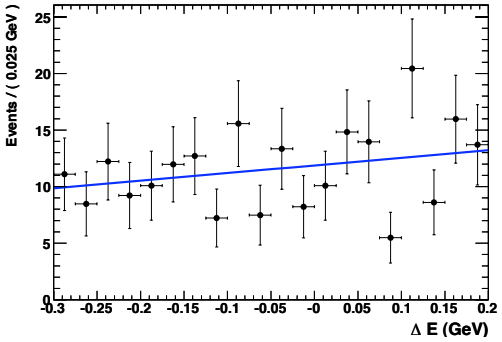
(b) $B \to K^*\gamma$ events where the candidate was incorrectly reconstructed (self cross-feed). The PDF is the 1D projection of the 2D histogram.



(c) $B \to X_s\gamma$ KN model events where the candidate was incorrectly reconstructed (mass cross-feed). The PDF is the 1D projection of the 2D histogram.

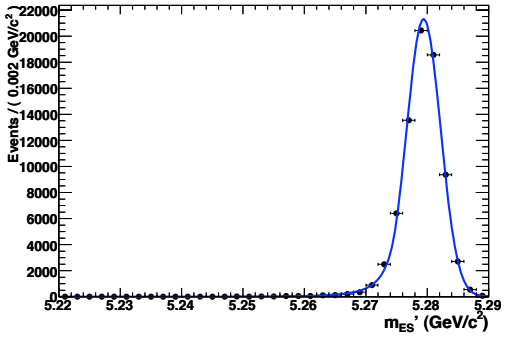


(d) Generic $B\overline{B}$ events modelled as an order-2 polynomial.

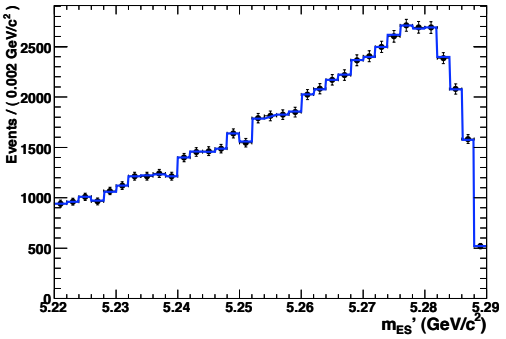


(e) Continuum events modelled as an order-1 polynomial.

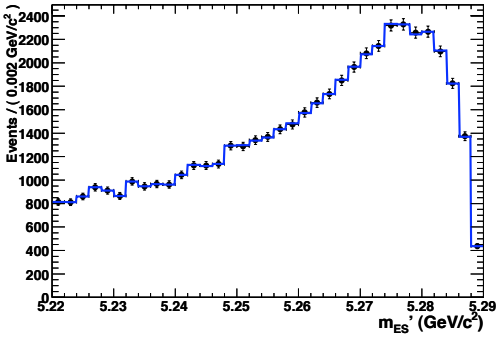
FIGURE 4.6: ΔE distribution PDF fits to MC data for $B \to X_s^L\gamma$ candidates. PDF distributions (solid blue line) are overlaid on the MC data whose error bars reflect the statistical uncertainty from the chosen binning of those data. In all cases the number of events indicated by the ordinate is arbitrarily determined from the MC data statistics.



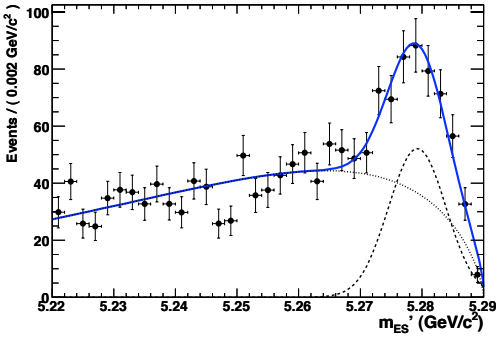
(a) $B \to X_s\gamma$ KN model events with a correctly reconstructed candidate (signal) modelled as a CB function.



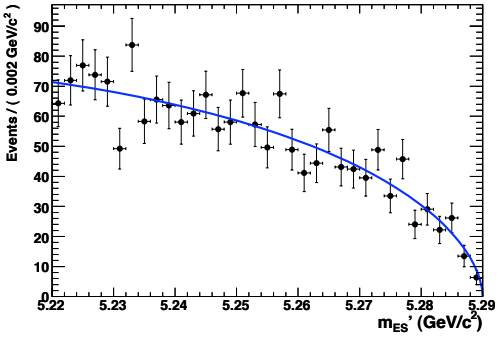
(b) $B \to X_s\gamma$ KN model events where the candidate was incorrectly reconstructed (self cross-feed). The PDF is the 1D projection of the 2D histogram.



(c) $B \to K^*\gamma$ events where the candidate was incorrectly reconstructed (mass cross-feed). The PDF is the 1D projection of the 2D histogram.

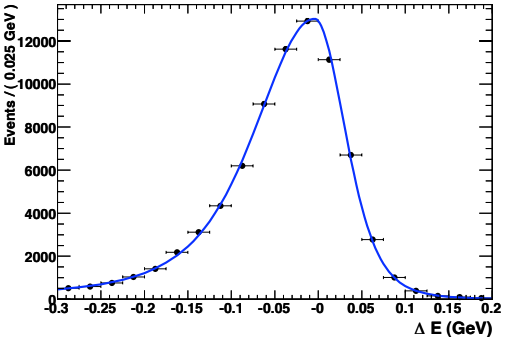


(d) Generic $B\overline{B}$ events modelled as the sum of a Gaussian (dashed) and Argus (dotted) function.

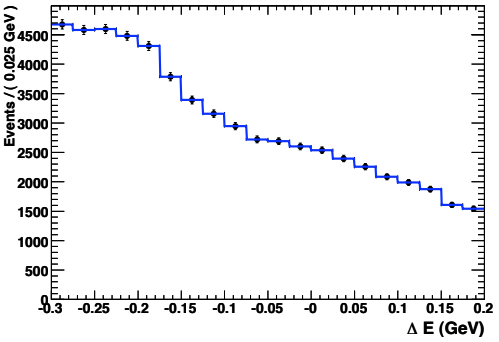


(e) Continuum events modelled as an Argus function.

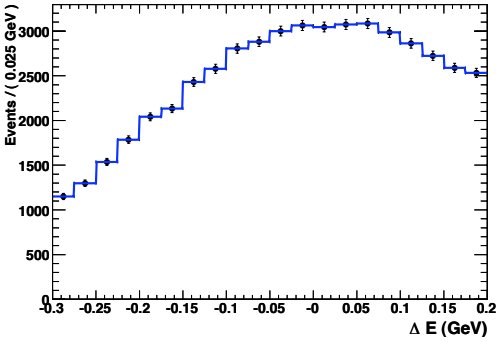
FIGURE 4.7: m'_{ES} distribution PDF fits to MC data for $B \to X_s^H\gamma$ candidates. PDF distributions (solid blue line) are overlaid on the MC data whose error bars reflect the statistical uncertainty from the chosen binning of those data. In all cases the number of events indicated by the ordinate is arbitrarily determined from the MC data statistics.



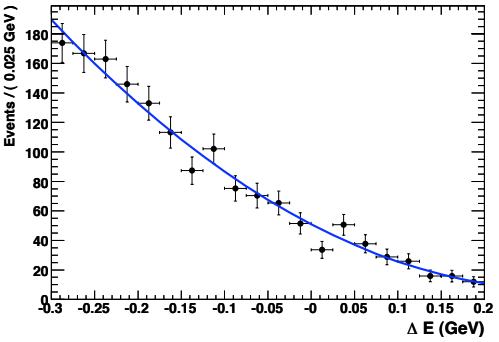
(a) $B \to X_s\gamma$ KN model events with a correctly reconstructed candidate (signal) modelled as a Cruijff function.



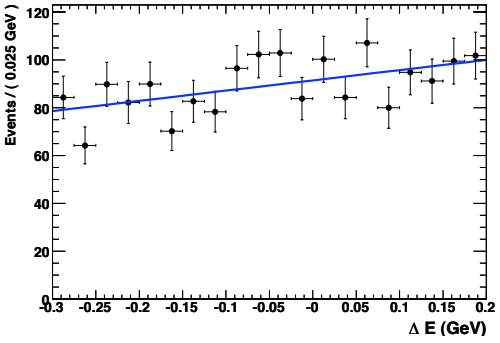
(b) $B \to X_s\gamma$ KN model events where the candidate was incorrectly reconstructed (self cross-feed). The PDF is the 1D projection of the 2D histogram.



(c) $B \to K^*\gamma$ events where the candidate was incorrectly reconstructed (mass cross-feed). The PDF is the 1D projection of the 2D histogram.



(d) Generic $B\overline{B}$ events modelled as an order-2 polynomial.



(e) Continuum events modelled as an order-1 polynomial.

FIGURE 4.8: ΔE distribution PDF fits to MC data for $B \to X_s^H\gamma$ candidates. PDF distributions (solid blue line) are overlaid on the MC data whose error bars reflect the statistical uncertainty from the chosen binning of those data. In all cases the number of events indicated by the ordinate is arbitrarily determined from the MC data statistics.

Hypothesis	Parameter	Low Mass Fit	High Mass Fit
Signal	μ (GeV)	-0.00307 ± 0.00037	-0.00445 ± 0.00068
	σ_L (GeV)	0.06875 ± 0.00044	0.06483 ± 0.00075
	σ_R (GeV)	0.03461 ± 0.00028	0.03412 ± 0.00052
	α_L	0.2125 ± 0.0023	0.2021 ± 0.0037
	α_L	0.1189 ± 0.0020	0.1148 ± 0.0036
Generic $B\overline{B}$	a_1 (GeV $^{-1}$)	-6.9 ± 1.7	-5.99 ± 0.29
	a_2 (GeV $^{-1}$)	8 ± 13	10.3 ± 2.1
Continuum	a_1 (GeV $^{-1}$)	0.57 ± 0.42	0.47 ± 0.16

TABLE 4.3: Fitted ΔE PDF parameters in MC data for each event hypothesis showing results for both low and high mass region $B \to X_s\gamma$ candidates.

4.4.3 $B \to X_d\gamma$ Candidate PDFs

Signal $B \to X_d\gamma$ events are modelled as a CB function in m'_{ES} and a Cruijff function in ΔE. Self cross-feed and mass cross-feed events are modelled as 2D histogram PDFs whose shapes are determined from MC data. PID cross-feed events will have very similar distributions to the signal but will be displaced in ΔE due to the difference in kaon and pion mass. They are therefore also modelled as a CB function in m'_{ES} and a Cruijff function in ΔE. X_s self cross-feed and mass cross-feed events are also modelled as 2D histogram PDFs with shapes determined from MC data. Generic $B\overline{B}$ events are modelled as the sum of a Gaussian and Argus function in m'_{ES}, with the Gaussian having relative normalisation, f, and a order-2 polynomial in ΔE. Finally continuum events are modelled as an Argus function in m'_{ES} and a order-1 polynomial in ΔE.

Figures 4.9 and 4.10 show the fits to MC data of m'_{ES} and ΔE respectively for $B \to X_d\gamma$ candidates in the low mass region and tables 4.4 and 4.5 list the respective estimator values and associated errors calculated from each PDF fit. The corresponding fits for $B \to X_d\gamma$ candidates in the high mass region are shown in figures 4.11 and 4.12 for m'_{ES} and ΔE respectively with the fitted estimator values and their errors also listed in tables 4.4 and 4.5.

4.5 Combined Fit Strategy

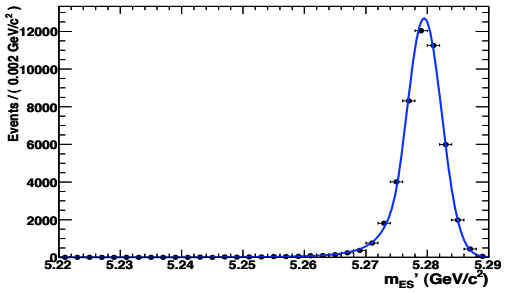
To extract the yield for each of the signal types from experimental data four corresponding EMLFs combining the relevant event hypotheses are required. This section describes the implementation of these fits and presents the results of test studies carried out to determine their stability and to check for any potential bias they may introduce.

Hypothesis	Parameter	Low Mass Fit	High Mass Fit
Signal	μ (GeV/c^2)	5.27949±0.00002	5.27939±0.00004
	σ (GeV/c^2)	0.002849±0.000014	0.002908±0.000034
	α	1.417±0.023	1.569±0.067
	n	3.85±0.15	3.21±0.31
X_s PID cross-feed	μ (GeV/c^2)	5.27918±0.00016	5.27931±0.00027
	σ (GeV/c^2)	0.00377±0.00014	0.00378±0.00021
	α	1.32±0.14	0.82±0.11
	n	2.67±0.55	14.1±9.3
Generic $B\overline{B}$	f	0.237±0.090	0.106±0.018
	μ (GeV/c^2)	5.2786±0.0011	5.28074±0.00049
	σ (GeV/c^2)	0.0034±0.0012	0.00290±0.00047
	ξ	−69±33	−49.7±5.7
	p	0.45±0.32	0.668±0.053
Continuum	ξ	−23±18	−19.9±2.9
	p	0.67±0.19	0.6673±0.0016

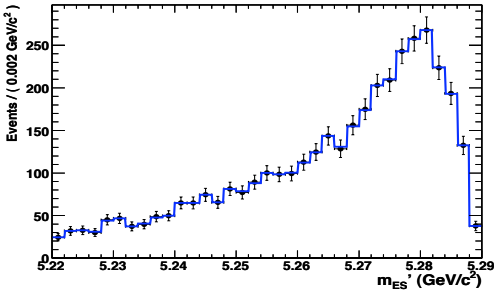
TABLE 4.4: Fitted m'_{ES} PDF parameters in MC data for each event hypothesis showing results for both low and high mass region $B \to X_d\gamma$ candidates.

Hypothesis	Parameter	Low Mass Fit	High Mass Fit
Signal	μ (GeV)	−0.00608±0.00090	−0.0070±0.0021
	σ_L (GeV)	0.0726±0.0011	0.0684±0.0026
	σ_R (GeV)	0.03673±0.00069	0.0352±0.0016
	α_L	0.2338±0.0060	0.237±0.014
	α_L	0.1061±0.0048	0.118±0.011
X_s PID cross-feed	μ (GeV)	−0.0836±0.0086	−0.086±0.010
	σ_L (GeV)	0.081±0.013	0.155±0.026
	σ_R (GeV)	0.0459±0.0061	0.0412±0.0068
	α_L	0.24±0.11	-0.64±0.34
	α_L	0.145±0.028	0.143±0.031
Generic $B\overline{B}$	a_1 (GeV $^{-1}$)	−9.5±1.6	−6.58±0.38
	a_2 (GeV $^{-1}$)	31±12	16.7±2.9
Continuum	a_1 (GeV $^{-1}$)	0.16±0.28	−0.29±0.14

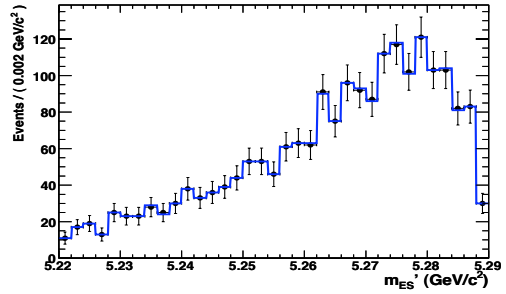
TABLE 4.5: Fitted ΔE PDF parameters in MC data for each event hypothesis showing results for both low and high mass region $B \to X_d\gamma$ candidates.



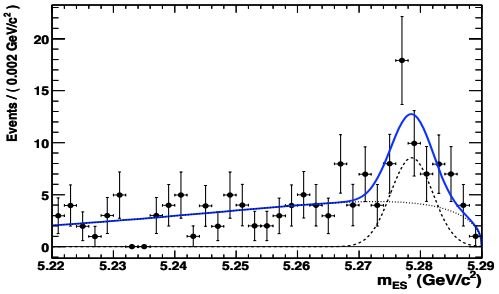
(a) $B \to \rho/\omega\gamma$ events with a correctly reconstructed candidate (signal) modelled as a CB function.



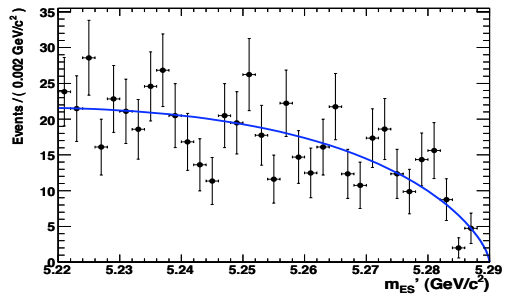
(b) $B \to \rho/\omega\gamma$ events where the candidate was incorrectly reconstructed (self cross-feed). The PDF is the 1D projection of the 2D histogram.



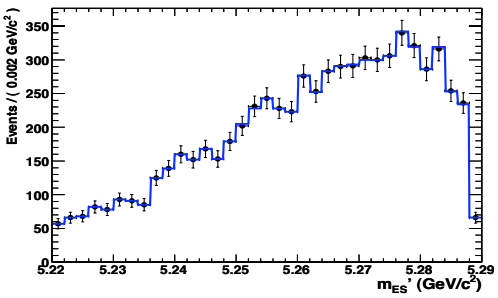
(c) $B \to X_d\gamma$ KN model events where the candidate was incorrectly reconstructed (mass cross-feed). The PDF is the 1D projection of the 2D histogram.



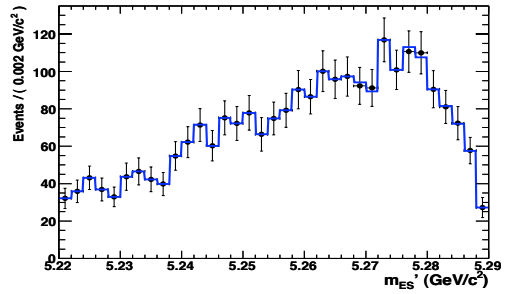
(d) Generic $B\overline{B}$ events modelled as the sum of a Gaussian (dashed) and Argus (dotted) function.



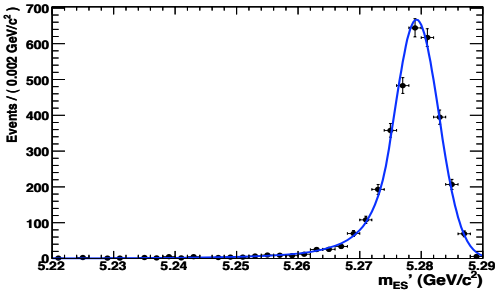
(e) Continuum events modelled as an Argus function.



(f) $B \to K^*\gamma$ events where the candidate was incorrectly reconstructed (X_s self cross-feed). The PDF is the 1D projection of the 2D histogram.

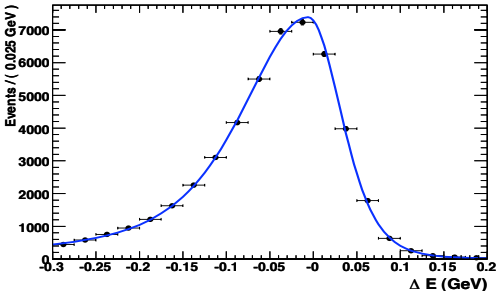


(g) $B \to X_s\gamma$ KN model events where the candidate was incorrectly reconstructed (X_s mass cross-feed). The PDF is the 1D projection of the 2D histogram.

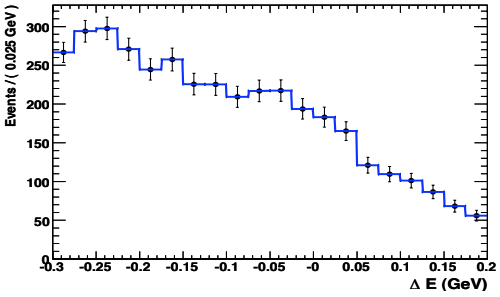


(h) $B \to X_s\gamma$ events with a correctly reconstructed candidate where the charged kaon is incorrectly identified as a charged pion (X_s PID cross-feed). Modelled as a CB function.

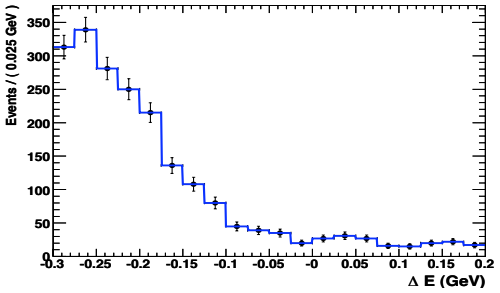
FIGURE 4.9: m'_{ES} distribution PDF fits to MC data for $B \to X_d^L\gamma$ candidates. PDF distributions (solid blue line) are overlaid on the MC data whose error bars reflect the statistical uncertainty from the chosen binning of those data. In all cases the number of events indicated by the ordinate is arbitrarily determined from the MC data statistics.



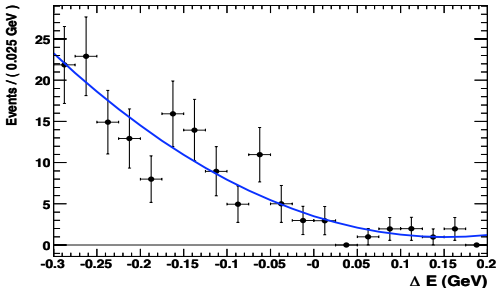
(a) $B \to \rho/\omega\gamma$ events with a correctly reconstructed candidate (signal) modelled as a Cruijff function.



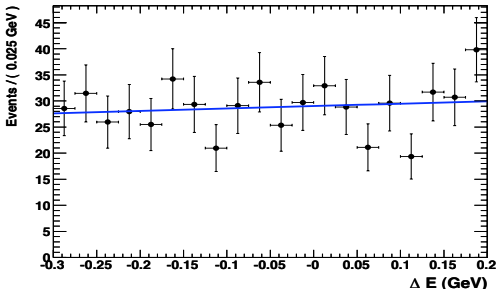
(b) $B \to \rho/\omega\gamma$ events where the candidate was incorrectly reconstructed (self cross-feed). The PDF is the 1D projection of the 2D histogram.



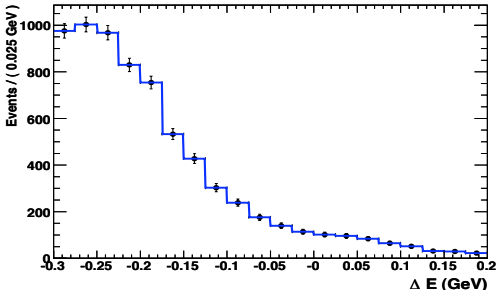
(c) $B \to X_d\gamma$ KN model events where the candidate was incorrectly reconstructed (mass cross-feed). The PDF is the 1D projection of the 2D histogram.



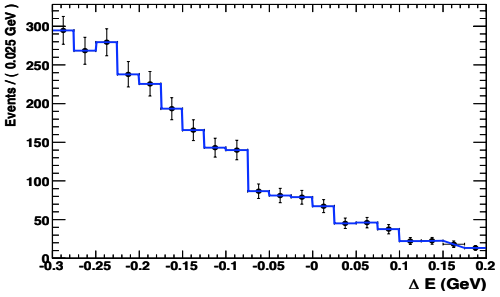
(d) Generic $B\overline{B}$ events modelled as an order-2 polynomial.



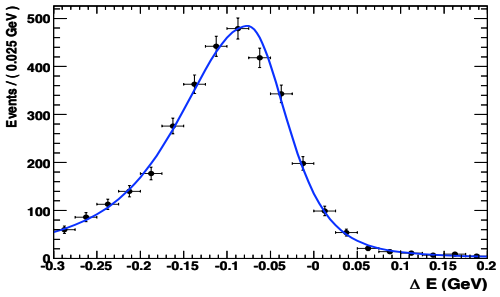
(e) Continuum events modelled as an order-1 polynomial.



(f) $B \to K^*\gamma$ events where the candidate was incorrectly reconstructed (X_s self cross-feed). The PDF is the 1D projection of the 2D histogram.

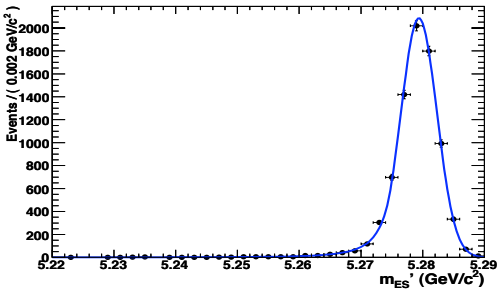


(g) $B \to X_s\gamma$ events where the candidate was incorrectly reconstructed (X_s mass cross-feed). The PDF is the 1D projection of the 2D histogram.

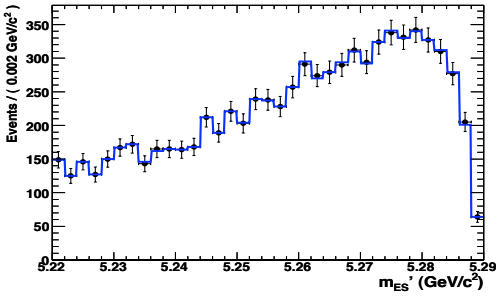


(h) $B \to X_s\gamma$ events with a correctly reconstructed candidate where the charged kaon is incorrectly identified as a charged pion (X_s PID cross-feed). Modelled as a Cruijff function.

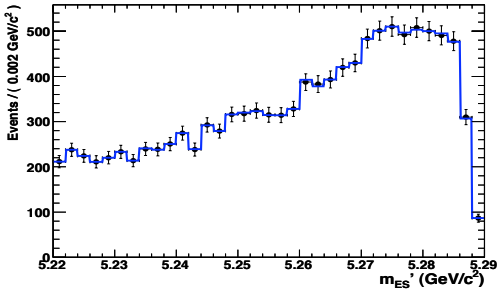
FIGURE 4.10: ΔE distribution PDF fits to MC data for $B \to X_d^L\gamma$ candidates. PDF distributions (solid blue line) are overlaid on the MC data whose error bars reflect the statistical uncertainty from the chosen binning of those data. In all cases the number of events indicated by the ordinate is arbitrarily determined from the MC data statistics.



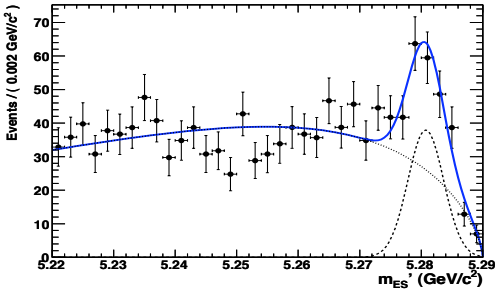
(a) $B \to X_d\gamma$ KN model events with a correctly reconstructed candidate (signal) modelled as a CB function.



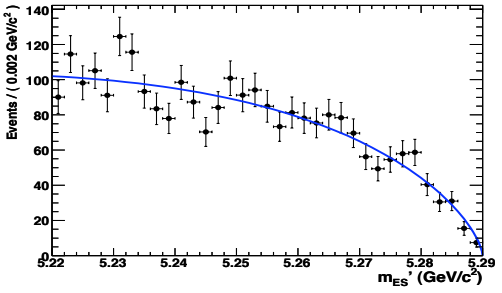
(b) $B \to X_d\gamma$ KN model events where the candidate was incorrectly reconstructed (self cross-feed). The PDF is the 1D projection of the 2D histogram.



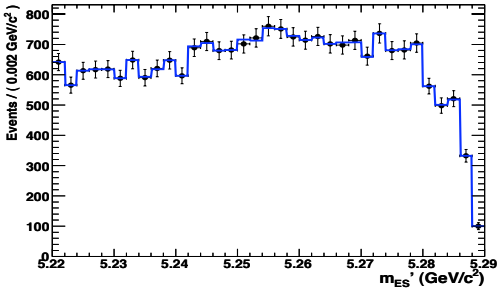
(c) $B \to \rho/\omega\gamma$ events where the candidate was incorrectly reconstructed (mass cross-feed). The PDF is the 1D projection of the 2D histogram.



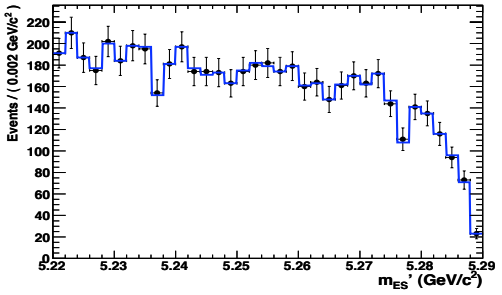
(d) Generic $B\overline{B}$ events modelled as the sum of a Gaussian (dashed) and Argus (dotted) function.



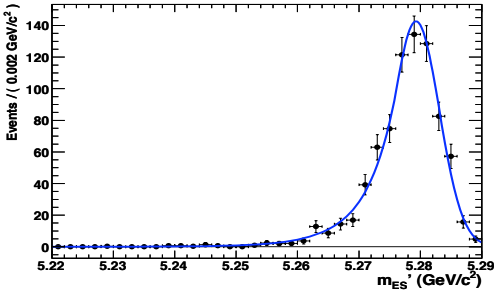
(e) Continuum events modelled as an Argus function.



(f) $B \to X_s\gamma$ KN model events where the candidate was incorrectly reconstructed (X_s self cross-feed). The PDF is the 1D projection of the 2D histogram.

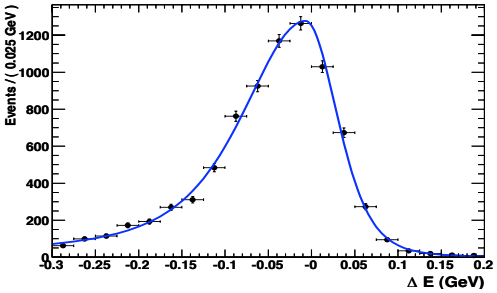


(g) $B \to K^*\gamma$ events where the candidate was incorrectly reconstructed (X_s mass cross-feed). The PDF is the 1D projection of the 2D histogram.

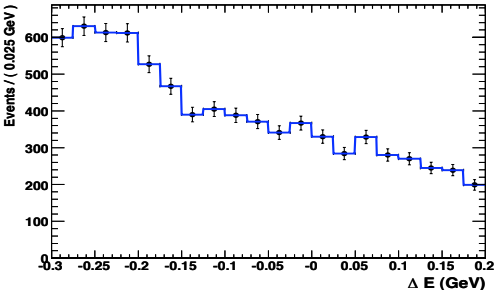


(h) $B \to X_s\gamma$ events with a correctly reconstructed candidate where the charged kaon is incorrectly identified as a charged pion (X_s PID cross-feed). Modelled as a CB function.

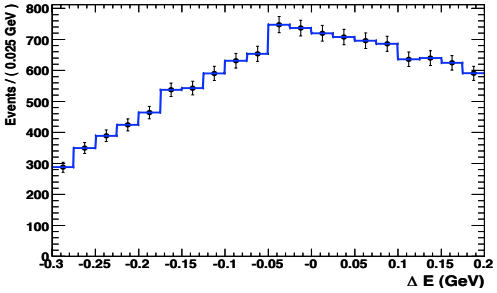
FIGURE 4.11: m'_{ES} distribution PDF fits to MC data for $B \to X_d^H\gamma$ candidates. PDF distributions (solid blue line) are overlaid on the MC data whose error bars reflect the statistical uncertainty from the chosen binning of those data. In all cases the number of events indicated by the ordinate is arbitrarily determined from the MC data statistics.



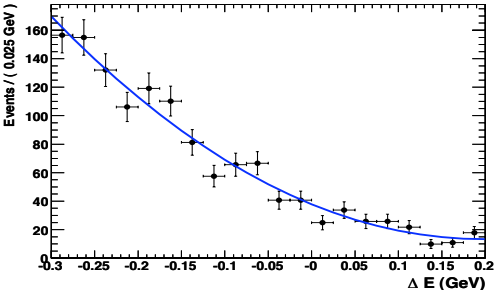
(a) $B \to X_d\gamma$ KN model events with a correctly reconstructed candidate (signal) modelled as a Cruijff function.



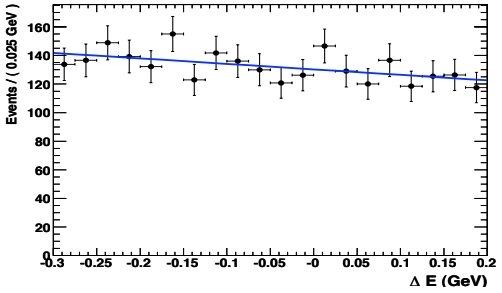
(b) $B \to X_d\gamma$ KN model events where the candidate was incorrectly reconstructed (self cross-feed). The PDF is the 1D projection of the 2D histogram.



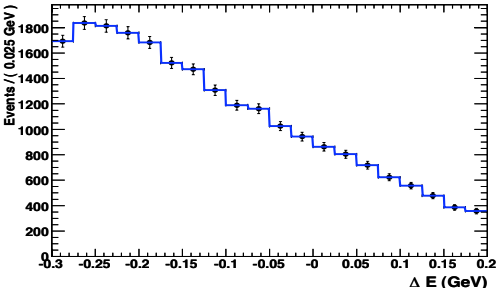
(c) $B \to \rho/\omega\gamma$ events where the candidate was incorrectly reconstructed (mass cross-feed). The PDF is the 1D projection of the 2D histogram.



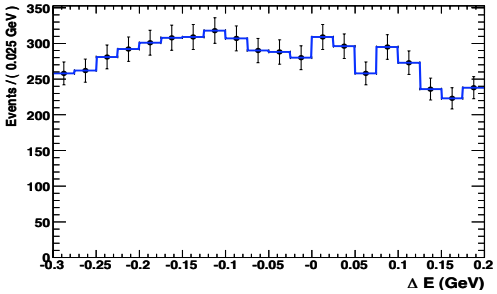
(d) Generic $B\overline{B}$ events modelled as an order-2 polynomial.



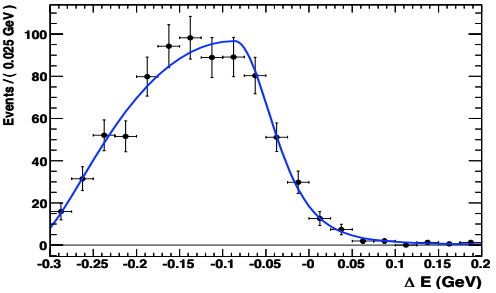
(e) Continuum events modelled as an order-1 polynomial.



(f) $B \to X_s\gamma$ KN model events where the candidate was incorrectly reconstructed (X_s self cross-feed). The PDF is the 1D projection of the 2D histogram.



(g) $B \to K^*\gamma$ events where the candidate was incorrectly reconstructed (X_s mass cross-feed). The PDF is the 1D projection of the 2D histogram.



(h) $B \to X_s\gamma$ events with a correctly reconstructed candidate where the charged kaon is incorrectly identified as a charged pion (X_s PID cross-feed). Modelled as a Cruijff function.

FIGURE 4.12: ΔE distribution PDF fits to MC data for $B \to X_d^H\gamma$ candidates. PDF distributions (solid blue line) are overlaid on the MC data whose error bars reflect the statistical uncertainty from the chosen binning of those data. In all cases the number of events indicated by the ordinate is arbitrarily determined from the MC data statistics.

4.5.1 Combined Fit Overview

The combined fit to experimental data follows the formalism of (4.2) with the EMLF mechanism resulting in the w_j weights becoming the mean number of events of each event hypothesis, n_j. The n_j and PDF parameters can either be fixed to pre-determined values or allowed to float in the fit to data. The signal yield, n_{sig}, must be floated as ultimately this is the measurement to be extracted from the experimental data. The choice of which other parameters to float is influenced by factors such as the data sample size, computational power, any correlation between those parameters and knowledge of inadequate modelling in MC data. Particularly important is the stability of the fit; floating too many variables may result in failure to find a stable minimum in l and therefore the fit will not converge. However, for each parameter which is fixed there is a potential of introducing a systematic bias in the signal yield.

Accumulated experience from many years of *BABAR* data analysis means there is generally reasonable agreement between simulated data distributions and their experimental data equivalents. However, for correctly reconstructed signal distributions, despite corrections during detector simulation to improve the modelling of energy leakage in the EMC, significant differences between widths in MC and experimental data for ΔE can still exist. The ΔE signal width parameters, σ_L and σ_R, are therefore floated when fitting for X_s candidates in both mass regions as the signal contributions in experimental data will be significant enough to sufficiently measure them. The corresponding values for these parameters in X_d candidate signal PDFs are then adjusted using the relative difference in X_s candidate MC and experimental data. All remaining signal MC parameters in ΔE and m'_{ES} are fixed to the values measured in MC data.

Cross-feed normalisations are fixed in fits to data. Previous analyses in the *BABAR* radiative penguin working group show that the ratio of signal to cross-feed is well modelled by the simulated data and should be preserved when fitting to experimental data [1, 30]. Each decay flavour is therefore fitted iteratively between mass regions with the relevant cross-feed yields corrected at each step to preserve their ratio relative to the new fitted signal yield. Iterations continue until the signal yields in both mass regions for that flavour converge. The $B \to X_s\gamma$ candidate data are fitted first and then the relevant fixed cross-feed yields of $B \to X_s\gamma$ events in $B \to X_d\gamma$ are modified to proportionally reflect the signal yields measured in the $B \to X_s\gamma$ data after convergence. Subsequently the $B \to X_d\gamma$ candidate data are fitted for, again iterating between the mass regions and updating the relevant $B \to X_d\gamma$ cross-feed yields at each step until convergence. In $B \to X_d\gamma$ fits the PID cross-feed PDF parameters are fixed to the values measured in MC data.

Generic $B\overline{B}$ events are generated using decay rate measurements from previous *BABAR* analyses and tend to show reasonable agreement between MC and experimental data. Therefore the normalisations and PDF parameters are fixed using MC data estimates in all fits. Continuum data distributions and normalisations are not so well modelled. In this analysis it is found that floating continuum PDF parameters in the low mass region fits result in them becoming unstable, therefore only the continuum yields are floated and the PDF parameters are fixed to their MC data values. In the high mass region the continuum Argus function shape parameter, ξ, and polynomial parameter, a_1 are also floated as this does not affect the fit stability.

Table 4.6 summarises which PDF parameters and yields are floated and which are fixed in fits to experimental data. The systematic uncertainties which arise from fixing fit components are estimated in chapter 5. Input values for variables floated in the combined fit to experimental data are determined from the MC data. PDF parameterisation variables are initially set to the values measured in MC data. The signal, cross-feed and continuum yields are initially set to the estimated data yields given in tables 3.22, 3.23 and where applicable table 4.1.

4.5.2 Studies to Test for Fit Stability and Bias

The fits to experimental data are tested for potential bias and stability with 'toy' MC studies. These repeat the fit numerous times with MC data where contributions from each event hypothesis are aggregated into an ensemble dataset intended to reflect the expected event distribution in real data. Such aggregate data can be either 'pure' or 'embedded'. Pure aggregate data is generated randomly, seeding events from the default PDF distributions and normalisations for each event hypothesis in the fit; the generated data therefore reflect those distributions. Embedded aggregate data comprise sub-samples of the fully simulated MC events for a given event hypothesis and normalisation. Fits to pure aggregate data determine the suitability of combining the PDFs used in a given fit, checking for any internal bias. Fits to embedded aggregate data will check the suitability of the PDFs used for those data as correlations in the fully simulated data may not be accounted for in the PDF distributions, hence leading to a bias in the embedded fits. When embedding events the sample size is Poisson distributed about the predetermined normalisation to reflect that an EMLF is to be used.

Toy fits to aggregate data determine the stability of the fit setup; if a significant number of these fits fail to converge then the fit setup will not be suitable for using on experimental data and the strategy must be reconsidered. Furthermore any bias which arises from floating a given parameter can be determined; the floated parameter will

Hypothesis	Category	Fixed	Floated
Signal	Yield		n_{sig}
	m_{ES} PDF	μ, σ, α, n	
	ΔE PDF ($B \to X_s\gamma$)	μ, α_L, α_R	σ_L, σ_R
	ΔE PDF ($B \to X_d\gamma$)	μ, α_L, α_R, σ_L, σ_R	
Self cross-feed	Yield	n_{scf} (corrected iteratively)	
Mass cross-feed	Yield	n_{mcf} (corrected iteratively)	
X_s PID cross-feed	Yield	n_{PID}^{Xs}	
($B \to X_d\gamma$ only)	m_{ES} PDF	μ, σ, α, n	
	ΔE PDF	μ, α_L, α_R, σ_L, σ_R	
X_s self cross-feed	Yield	n_{scf}^{Xs}	
($B \to X_d\gamma$ only)			
X_s mass cross-feed	Yield	n_{mcf}^{Xs}	
($B \to X_d\gamma$ only)			
Generic $B\overline{B}$	Yield	n_{GenB}	
	m_{ES} PDF	ξ, p, f, μ, σ	
	ΔE PDF	a_1, a_2	
Continuum	Yield		n_{cont}
	m_{ES} PDF ($B \to X^L\gamma$)	ξ, p	
	ΔE PDF ($B \to X^L\gamma$)	a_1	
	m_{ES} PDF ($B \to X^H\gamma$)	p	ξ
	ΔE PDF ($B \to X^H\gamma$)		a_1

TABLE 4.6: Summary of fixed and floated event class yields and PDF parameters in fits to experimental data.

have a known value when the aggregate data is generated, α_{gen}. The value returned from the toy fit, α_{fit}, will then have an associated 'pull' which measures the difference in the generated and fitted values scaled by the error on that parameter calculated from the fit, σ_α; explicitly the pull is defined as $(\alpha_{gen} - \alpha_{fit})/\sigma_\alpha$. For an unbiased fit the pull distribution accumulated from numerous identical fits to different aggregate data samples will approach a Gaussian distribution centred on zero with unit width. Any departure from this distribution indicates internal bias in the fit which systematically biases the parameter under consideration. Such biases can result from an ill-defined fit strategy (for example fixing one parameter which is highly correlated with a floated parameter) or simply due to too few statistics for a given event hypothesis. This bias must be removed by revising the fit strategy or corrected for after the fit to data. In this analysis only potential bias on the signal yield is corrected for.

For each of the four fits to be carried out on experimental data, 500 toy fits to pure aggregate data are performed to check the suitability of combining the different event

hypothesis PDFs in the fit. Additionally 500 toy fits embedding signal and cross-feed simulated data are performed to investigate potential bias in the floated yields due to these data. The resulting pull distribution histograms are fitted with a Gaussian function using a χ^2 minimisation technique [63].

The toy study framework does not allow the iteration between mass regions to correct cross-feed components to be studied, so the cross-feed yields are additionally varied as part of the systematic studies described in chapter 5. Furthermore the framework does not persist the weighted corrections made to $B \to X_s\gamma$ phase space decays described in section 4.2 so all embedded studies are carried out with PDF distributions and normalisations determined from unweighted MC data. It is not believed that weighting the MC data will have a significant effect on the signal yield bias, so the unweighted embedded studies are still considered a robust investigation of fit bias. The pure toy studies do use the weighted MC data distributions and normalisations.

4.5.2.1 X_s Candidate Fits

Figure 4.13 shows the signal yield pull histograms and fitted Gaussian distributions from aggregate MC data toy studies for fits to low and high mass $B \to X_s\gamma$ candidates. The embedded toy studies require the signal, self cross-feed and mass cross-feed components to be sampled from fully simulated MC data and the remaining event hypotheses generated as pure aggregate MC data. Tables 4.7 and 4.8 list the fitted Gaussian parameters, μ and σ, to pull histograms for all components floated in the respective low and high mass fits. Also shown are the average correlation coefficients, ρ_{sig}, between each floated variable and the signal yield calculated from the embedded toy studies. In all cases the signal yield distributions are deemed acceptable and the average pull on the embedded signal yield distributions will be used to correct the signal yield obtained from fits to experimental data. Pulls calculated from each floated variable in the embedded study are used to estimate systematic uncertainties due to fit bias in section 5.1.3.4. Low correlation coefficients mean that the relatively large pulls observed for some signal PDF parameters are expected to have a relatively small effect on n_{sig} and so are not a significant concern. All fits performed in these studies converged successfully.

4.5.2.2 X_d Candidates

Figure 4.14 shows the signal yield pull histograms and fitted Gaussian distributions from aggregate MC data toy studies for the fits to low and high mass $B \to X_d\gamma$ candidates. Two distinct embedded toy studies are carried out. Initially the signal, X_d self cross-feed and X_d mass cross-feed components are sampled from fully simulated MC data

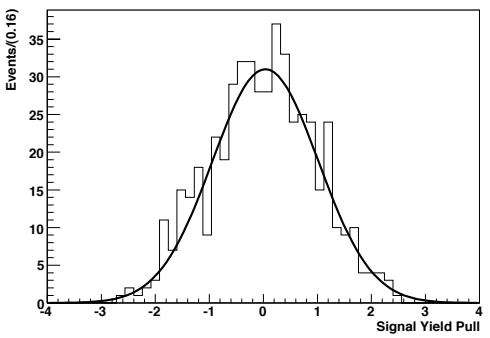
Parameter		Pure Study	Embedded Study	ρ_{sig}
n_{sig}	μ	0.038±0.047	−0.127±0.048	1.000
	σ	0.981±0.037	0.999±0.037	
n_{cont}	μ	−0.106±0.047	−0.092±0.047	-0.069
	σ	0.978±0.044	0.998±0.037	
σ_L	μ	−0.003±0.048	−0.461±0.054	0.098
	σ	0.983±0.035	1.100±0.043	
σ_R	μ	0.070±0.051	−0.282±0.045	0.042
	σ	1.012±0.040	0.964±0.034	

TABLE 4.7: Fitted Gaussian parameters to pull distributions from toy MC studies for all floated parameters in the low mass $B \to X_s\gamma$ fit. Also shown are the average correlation coefficients between each floated variable and n_{sig} for the embedded toy studies.

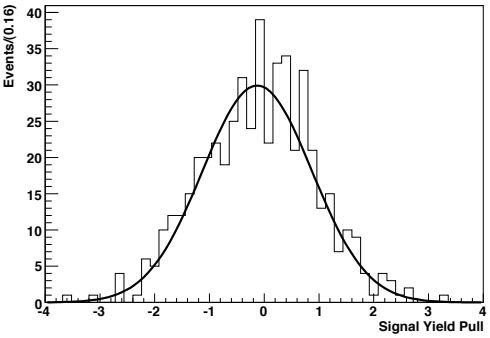
Parameter		Pure Study	Embedded Study	ρ_{sig}
n_{sig}	μ	−0.050±0.044	0.017±0.045	1.000
	σ	0.924±0.039	0.927±0.036	
n_{cont}	μ	−0.030±0.044	0.108±0.045	-0.202
	σ	0.947±0.034	0.930±0.035	
σ_L	μ	−0.077±0.049	−0.834±0.046	0.233
	σ	1.014±0.039	0.954±0.034	
σ_R	μ	−0.097±0.043	−0.326±0.047	0.167
	σ	0.926±0.035	0.970±0.038	
ξ	μ	−0.047±0.047	0.002±0.043	0.251
	σ	0.968±0.036	0.895±0.036	
a_1	μ	−0.003±0.046	−0.135±0.050	0.048
	σ	0.943±0.037	1.042±0.041	

TABLE 4.8: Fitted Gaussian parameters to pull distributions from toy MC studies for all floated parameters in the high mass $B \to X_s\gamma$ fit. Also shown are the average correlation coefficients between each floated variable and n_{sig} for the embedded toy studies.

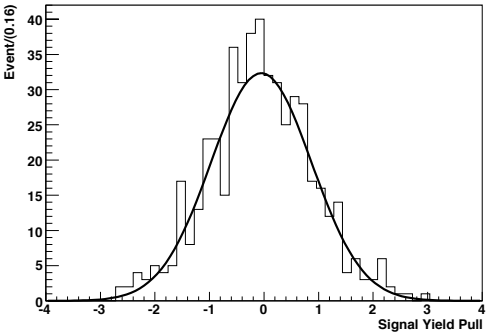
and the remaining event hypotheses are generated as pure aggregate MC data. Finally the signal, X_d self cross-feed, X_d mass cross-feed, X_s PID cross-feed, X_s self cross-feed and X_s mass cross-feed components are sampled from fully simulated MC data and the remaining event hypotheses are generated as pure aggregate MC data. Tables 4.9 and 4.10 list the fitted Gaussian parameters, μ and σ, to pull histograms for all components floated in the respective low and high mass regions for these toy studies. Also shown are the average correlation coefficients between each floated variable and n_{sig} for the embedded toy study where both X_d and X_s MC data are embedded. In all cases the distributions are deemed acceptable and the average pull on the signal yield distribution with both X_d and X_s cross-feed events embedded will be used to correct the signal yield



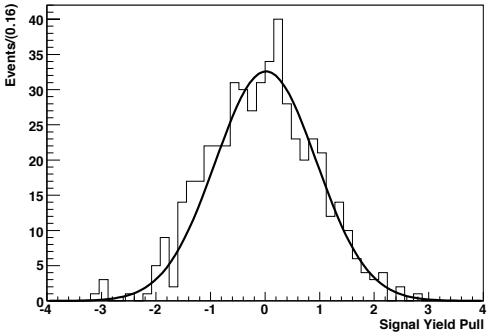
(a) Low mass pure toy signal yield pull.



(b) Low mass embedded signal and X_s cross-feed toy signal yield pull.



(c) High mass pure toy signal yield pull.



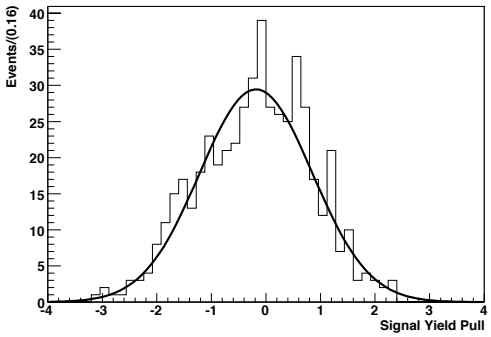
(d) High mass embedded signal and X_s cross-feed toy signal yield pull.

FIGURE 4.13: Toy study signal yield pull histograms and fitted Gaussian distributions for $B \to X_s\gamma$ candidate fits to experimental data. Each study consists of 500 toy fits.

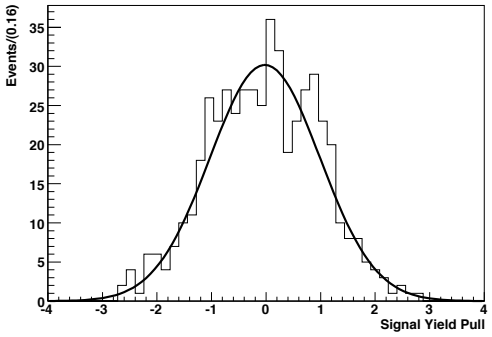
Parameter		Pure Study	Embedded Study	Embedded X_s Study	ρ_{sig}
n_{sig}	μ	-0.182 ± 0.049	-0.021 ± 0.048	0.109 ± 0.045	1.000
	σ	1.033 ± 0.038	1.010 ± 0.035	0.940 ± 0.041	
n_{cont}	μ	0.051 ± 0.049	0.009 ± 0.048	-0.025 ± 0.045	-0.151
	σ	1.051 ± 0.042	1.001 ± 0.039	0.947 ± 0.036	

TABLE 4.9: Fitted Gaussian parameters to pull distributions from toy MC studies for all floated parameters in the low mass $B \to X_d\gamma$ fit. Also shown are the average correlation coefficients between each floated variable and n_{sig} for the embedded toy study where both X_d and X_s MC data are embedded.

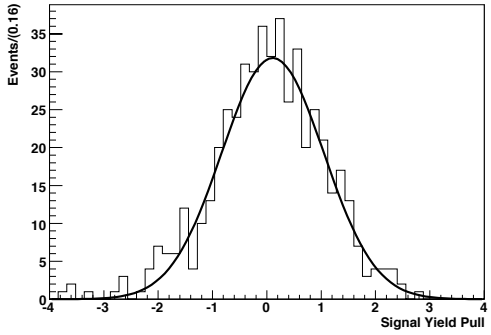
obtained from fits to experimental data. Pulls from each floated variable in this full embedded study are used to evaluate a fit bias systematic uncertainty in section 5.1.3.4. All fits performed in these studies converged successfully.



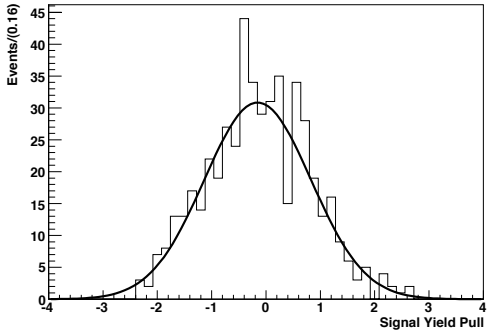
(a) Low mass pure toy signal yield pull.



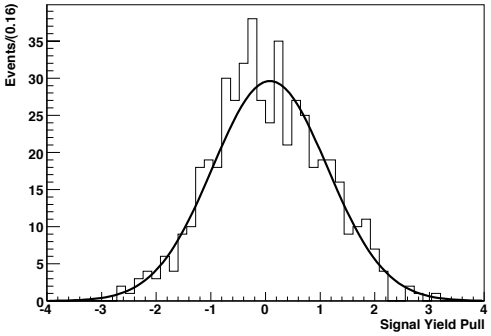
(b) Low mass embedded signal and X_d cross-feed toy signal yield pull.



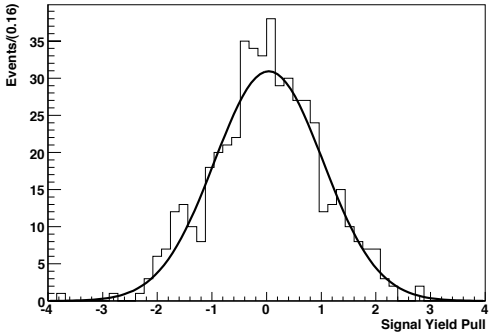
(c) Low mass embedded signal, X_d cross-feed and X_s cross-feed toy signal yield pull.



(d) High mass pure toy signal yield pull.



(e) High mass embedded signal and X_d cross-feed toy signal yield pull.



(f) High mass embedded signal, X_d cross-feed and X_s cross-feed toy signal yield pull.

FIGURE 4.14: Toy study signal yield pull histograms and fitted Gaussian distributions for $B \to X_d\gamma$ candidate fits to experimental data. Each study consists of 500 toy fits.

Parameter		Pure Study	Embedded Study	Embedded X_s Study	ρ_{sig}
n_{sig}	μ	-0.153 ± 0.048	0.082 ± 0.052	0.041 ± 0.048	1.000
	σ	0.983 ± 0.041	1.051 ± 0.043	0.995 ± 0.043	
n_{cont}	μ	0.036 ± 0.047	-0.068 ± 0.045	0.050 ± 0.043	-0.222
	σ	0.997 ± 0.038	0.937 ± 0.037	0.893 ± 0.036	
ξ	μ	-0.104 ± 0.050	-0.027 ± 0.046	-0.097 ± 0.046	0.326
	σ	1.031 ± 0.043	0.986 ± 0.037	0.979 ± 0.035	
a_1	μ	0.057 ± 0.044	0.028 ± 0.046	-0.021 ± 0.050	0.026
	σ	0.907 ± 0.035	0.951 ± 0.036	1.031 ± 0.043	

TABLE 4.10: Fitted Gaussian parameters to pull distributions from toy MC studies for all floated parameters in the high mass $B \to X_d\gamma$ fit. Also shown are the average correlation coefficients between each floated variable and n_{sig} for the embedded toy study where both X_d and X_s MC data are embedded.

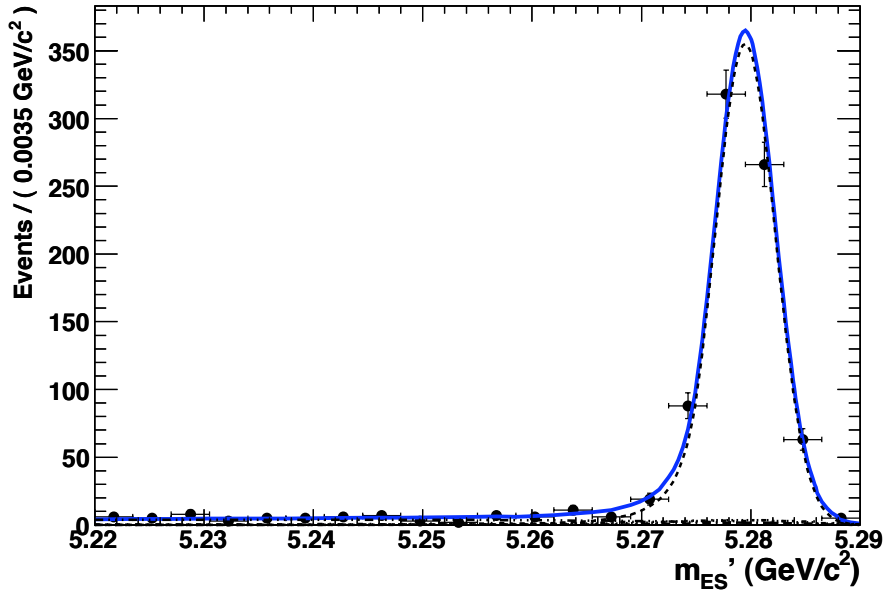
4.6 Fits To Experimental Data

The fits described above are applied to experimental data passing event selection cuts in each of the four event reconstruction categories. The results of these fits and their physical interpretation are presented here.

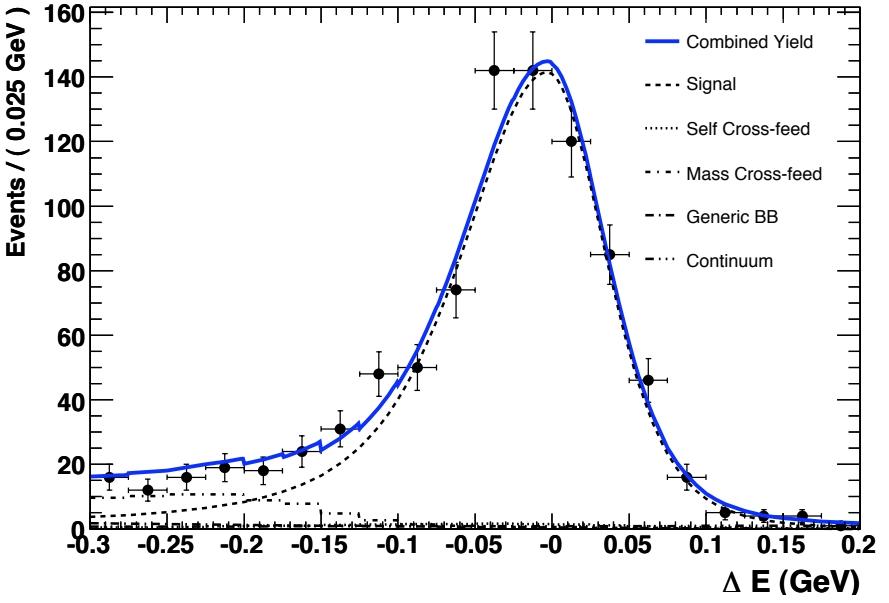
4.6.1 X_s Candidate Fit Results

The $B \to X_s\gamma$ fits to experimental data converge after four complete iterations between the mass regions to correct cross-feed yields. Tables 4.11 and 4.12 list the initial, final and bias corrected values with associated errors for all variables floated in the fits to data in the low and high mass regions respectively. Bias corrections are applied by multiplying the average pull from the relevant toy study by the error returned from the experimental fit and then subtracting the resulting product from the fitted value for that variable. The tables also show the revised cross-feed yields after iteration. Figure 4.15 shows projection plots of the combined fit model and key contributing backgrounds for m'_{ES} and ΔE in the low mass region. The corresponding projection plots for the high mass region are shown in figure 4.16. The m'_{ES} projection plots are produced after applying a cut of $-0.15 \leq \Delta E \leq 0.1\,\text{GeV}$ and the ΔE projection plots are produced after applying a cut of $m'_{ES} \geq 5.27\,\text{GeV}/c^2$. All projection plots show the fit model is in good agreement with the experimental data in both the high and low mass regions.

The low mass signal yield from experimental data is in reasonable agreement with the expected signal yield. The high mass signal yield is lower than the corresponding estimated value from MC studies. However, this does not account for systematic effects which are estimated in chapter 5 and so is not thought to be significant.

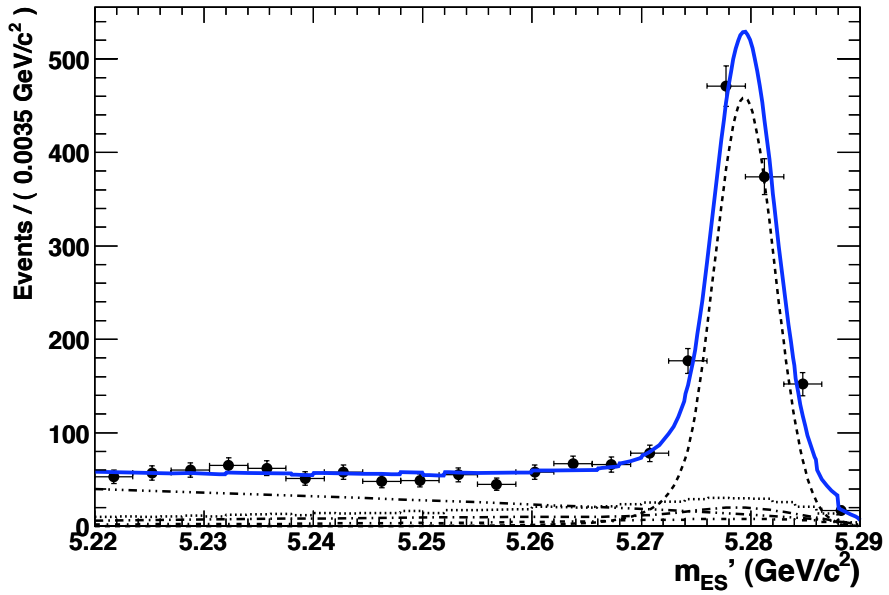


(a) m'_{ES} projection plot

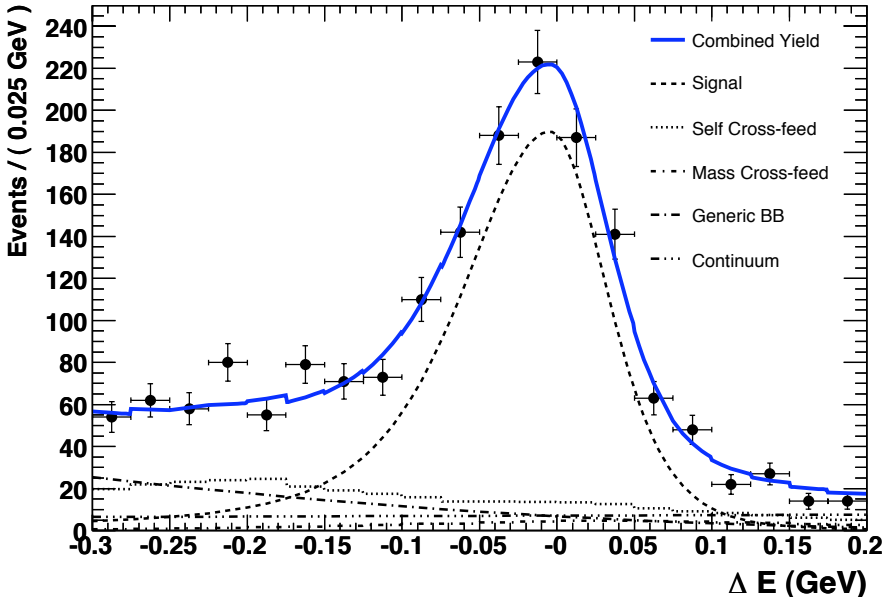


(b) ΔE projection plot

FIGURE 4.15: Projection plots of m'_{ES} and ΔE for $B \to X_s^L \gamma$ candidate fit to experimental data. The legend is applicable to both plots and indicates the relative contribution from each event hypothesis to the overall combined yield in experimental data.



(a) m'_{ES} projection plot



(b) ΔE projection plot

FIGURE 4.16: Projection plots of m'_{ES} and ΔE for $B \to X_s^H \gamma$ candidate fit to experimental data. The legend is applicable to both plots and indicates the relative contribution from each event hypothesis to the overall combined yield in experimental data.

Parameter	Initial Value	Final Value	Fit Bias Corrected Value
n_{sig}	801	770±29	774±29
n_{cont}	144	118±14	119±14
σ_L	0.0688±0.0004	0.0523±0.0029	0.0536±0.0029
σ_R	0.0346±0.0003	0.0365±0.0018	0.0370±0.0018
n_{scf}	32	31	-
n_{mcf}	155	149	-

TABLE 4.11: Result of fit to experimental data for $B \to X_s^L\gamma$ candidates showing initial, final and bias-corrected values for variables floated in the fit as well as the corrected cross-feed yields after iteration.

Parameter	Initial Value	Final Value	Fit Bias Corrected Value
n_{sig}	1112	1016±39	1015±39
n_{cont}	1113	996±51	990±51
σ_L	0.0648±0.0008	0.0532±0.0032	0.0559±0.0032
σ_R	0.0341±0.0005	0.0355±0.0019	0.0361±0.0019
ξ	−9.3±9.3	4.6±8.4	4.6±8.4
a_1	0.47±0.16	0.31±0.32	0.35±0.32
n_{scf}	879	803	-
n_{mcf}	155	149	-

TABLE 4.12: Result of fit to experimental data for $B \to X_s^H\gamma$ candidates showing initial, final and bias-corrected values for variables floated in the fit as well as the corrected cross-feed yields after iteration.

4.6.2 X_d Candidate Fit Results

Before fitting for $B \to X_d\gamma$ candidates in experimental data the expected X_s cross-feed yields and ΔE signal distribution widths are corrected using the results from fits to $B \to X_s\gamma$ candidates in experimental data. Table 4.13 lists the modified X_s cross-feed component yields. They are corrected by applying the ratio of corresponding measured and estimated signal yield obtained from the $B \to X_s\gamma$ candidate fits. Table 4.14 lists the modified ΔE signal width parameters, σ'_{Xd}, given by,

$$\sigma'_{Xd} = \sigma_{Xd}^{MC}\frac{\sigma_{Xs}^{data}}{\sigma_{Xs}^{MC}}, \tag{4.14}$$

where σ_{Xd}^{MC} is the corresponding fitted width in $B \to X_d\gamma$ candidate signal MC, σ_{Xs}^{data} is the bias corrected width measured in $B \to X_s\gamma$ candidate experimental data and σ_{Xs}^{MC} is the width from fitting $B \to X_s\gamma$ candidate signal MC data. The associated error for σ'_{Xd} is obtained from,

$$\epsilon = \sigma'_{Xd}\frac{\epsilon_{Xs}^{data}}{\sigma_{Xs}^{data}}, \tag{4.15}$$

(a) Low mass $B \to X_d\gamma$ candidate yields.

Yield	Initial	Corrected
n_{PID}^{Xs}	13	13
n_{scf}^{Xs}	22	21
n_{mcf}^{Xs}	36	33

(b) High mass $B \to X_d\gamma$ candidate yields.

Yield	Initial	Corrected
n_{PID}^{Xs}	11	10
n_{scf}^{Xs}	324	295
n_{mcf}^{Xs}	18	17

TABLE 4.13: Corrections of $B \to X_s\gamma$ cross-feed component yields in (a) the low mass and (b) the high mass fit to $B \to X_d\gamma$ candidates after fitting for $B \to X_s\gamma$ candidates in experimental data.

(a) Low mass $B \to X_d\gamma$ fit.

Width	Initial	Corrected
σ_L	0.0726±0.0011	0.0566±0.0031
σ_R	0.0367±0.0007	0.0392±0.0019

(b) High mass $B \to X_d\gamma$ fit.

Width	Initial	Corrected
σ_L	0.0684±0.0026	0.0590±0.0034
σ_R	0.0352±0.0016	0.0373±0.0020

TABLE 4.14: Corrections to ΔE signal widths in (a) the low mass and (b) the high mass fit to $B \to X_d\gamma$ candidates after fitting for $B \to X_s\gamma$ candidates in experimental data.

where ϵ_{Xs}^{data} is the associated error from the fit to $B \to X_s\gamma$ candidate experimental data.

The $B \to X_d\gamma$ fits to experimental data converge after three complete iterations between the mass regions to correct cross-feed yields. Tables 4.15 and 4.16 list the initial, final and bias corrected values with associated errors for all variables floated in the fits to data in the low and high mass regions respectively. Also shown are the revised cross-feed yields after iteration. Figure 4.17 shows projection plots of the combined fit model and key contributing backgrounds for m'_{ES} and ΔE in the low mass region. The corresponding projection plots for the high mass region are shown in figure 4.16. The m'_{ES} projection plots are produced after applying a cut of $-0.15 \leq \Delta E \leq 0.1\,\text{GeV}$ and the ΔE projection plots are produced after applying a cut of $m'_{ES} \geq 5.27\,\text{GeV}/c^2$. All projection plots show the fit model is in good agreement with the experimental data in both the high and low mass regions.

The low mass signal yield in experimental data is in good agreement with the expected yield. For the high mass signal yield an enhancement compared to MC data estimates is observed. However, these estimates use an order of magnitude calculation for the expected inclusive $B \to X_d\gamma$ BF. It follows that the initial and expected values should

Parameter	Initial Value	Final Value	Fit Bias Corrected Value
n_{sig}	35	34±9	33±9
n_{cont}	404	366±22	367±22
n_{scf}	3	3	-
n_{mcf}	7	12	-

TABLE 4.15: Result of fit to experimental data for $B \to X_d^L\gamma$ candidates showing initial, final and bias-corrected values for variables floated in the fit as well as the corrected cross-feed yields after iteration.

Parameter	Initial Value	Final Value	Fit Bias Corrected Value
n_{sig}	37	60±15	59±15
n_{cont}	1722	1450±49	1448±49
ξ	-19.8±2.9	-5.0±5.3	-4.5±5.3
a_1	-0.29±0.14	-0.12±0.22	-0.12±0.22
n_{scf}	37	60	-
n_{mcf}	8	8	-

TABLE 4.16: Result of fit to experimental data for $B \to X_d^H\gamma$ candidates showing initial, final and bias-corrected values for variables floated in the fit as well as the corrected cross-feed yields after iteration.

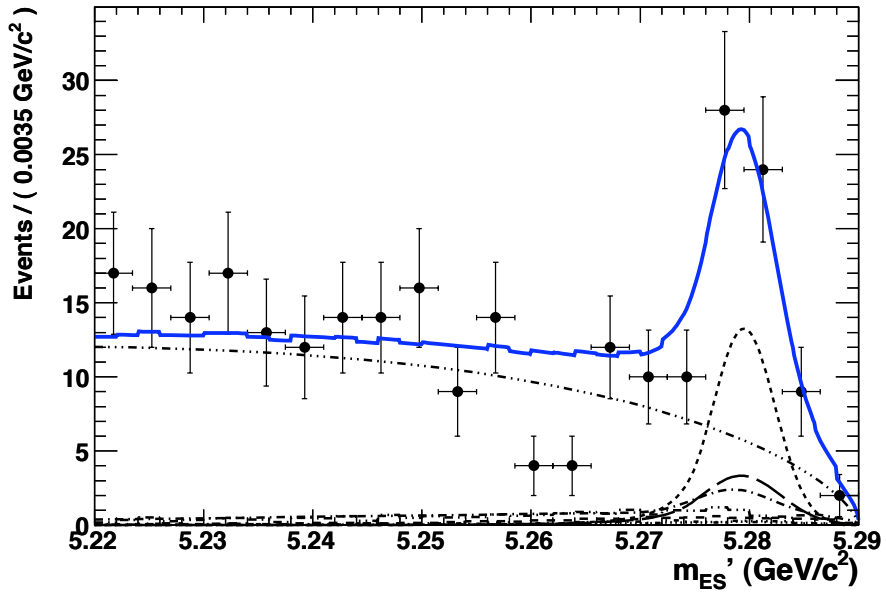
therefore only be expected to agree in order of magnitude and this is indeed the case. The statistical significance of this measurement does improve significantly on the previous *BABAR* measurement [1] and represents the first evidence for $B \to X_d\gamma$ transitions in the hadronic mass region above the ρ and ω resonances.

4.6.3 Quality of Fits to Experimental Data

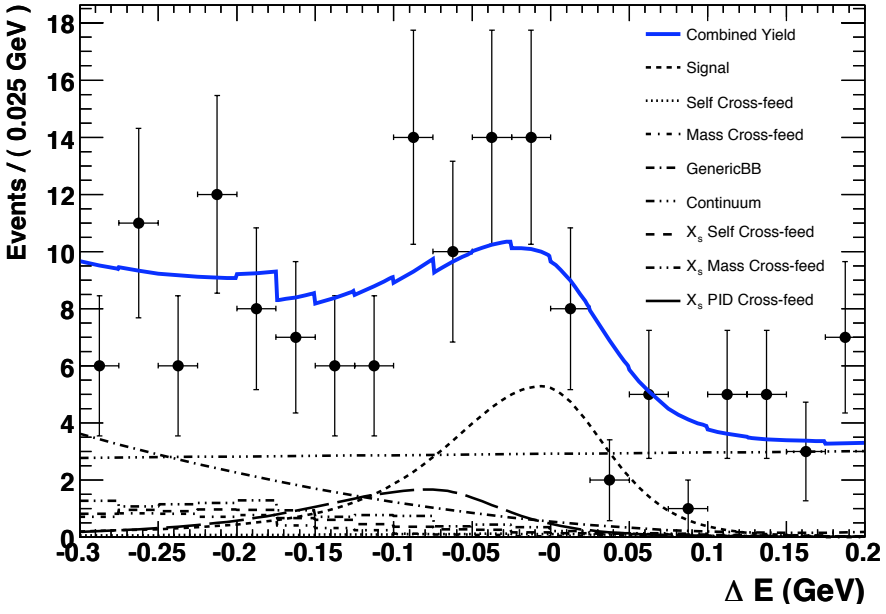
Agreement between each combined yield PDF distribution and corresponding experimental data is evaluated numerically using a chi square goodness of fit test [63, 64]. Each 2D combined yield PDF is integrated numerically in a grid of 20 bins in ΔE and 20 bins[5] in m'_{ES} for this calculation. The integrated grid is then projected in each variable with a cut placed on the signal region of the orthogonal variable, thus emulating the projected PDF distributions shown above as binned histograms. The corresponding experimental data is projected identically for each variable. The chi square statistic comparing the PDF and experimental data histograms is given by,

$$\chi^2 = \sum_k \frac{(n^k_{PDF} - n^k_{Data})^2}{\sigma^2_{Data}}, \tag{4.16}$$

[5]The binning in m'_{ES} is reduced from 35 bin to 20 bins for this calculation to ensure sufficient statistics per bin in the projection of this variable. In particular in the $B \to X_d^L\gamma$ fit has relatively few experimental data events.

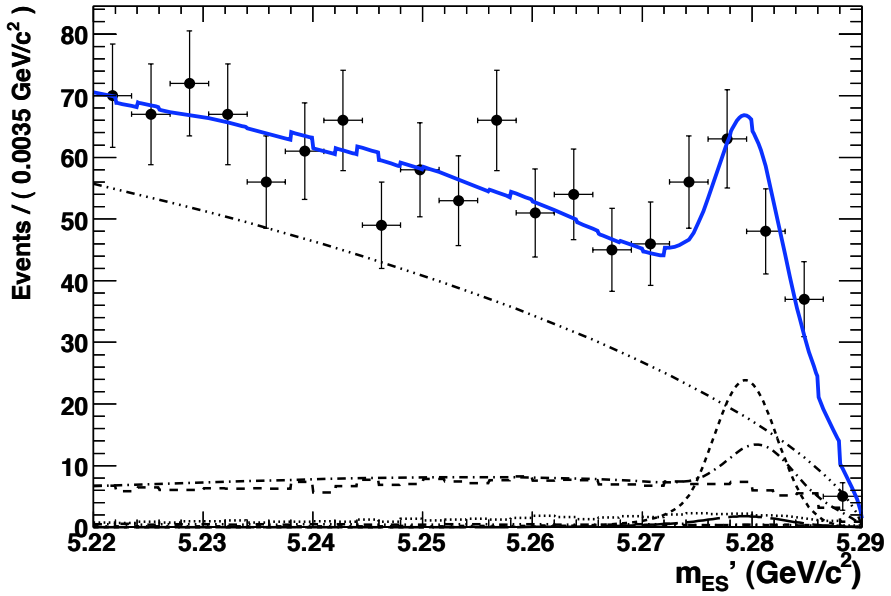


(a) m'_{ES} projection plot

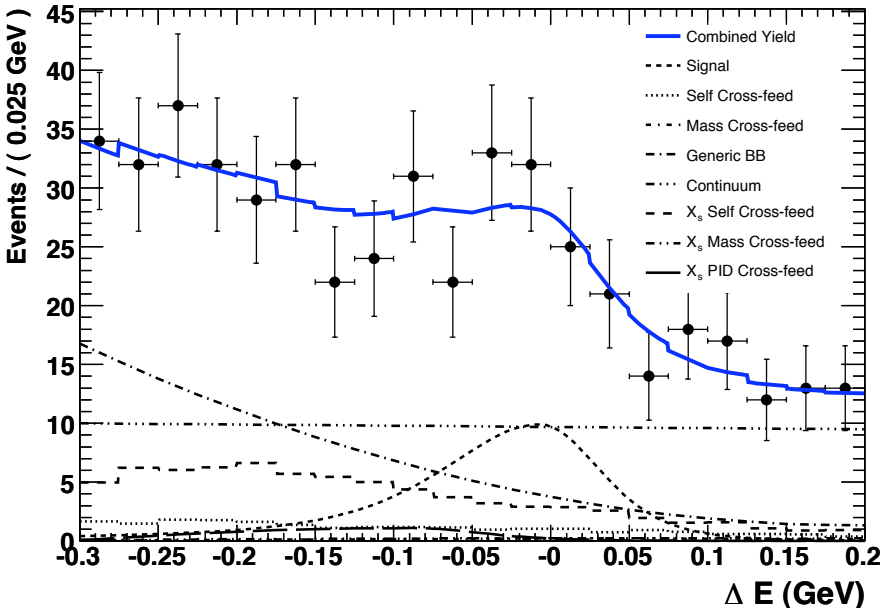


(b) ΔE projection plot

FIGURE 4.17: Projection plots of m'_{ES} and ΔE for $B \to X_d^L \gamma$ candidate fit to experimental data. The legend is applicable to both plots and indicates the relative contribution from each event hypothesis to the overall combined yield in experimental data.



(a) m'_{ES} projection plot



(b) ΔE projection plot

FIGURE 4.18: Projection plots of m'_{ES} and ΔE for $B \to X_d^H \gamma$ candidate fit to experimental data. The legend is applicable to both plots and indicates the relative contribution from each event hypothesis to the overall combined yield in experimental data.

Fit	Variable	χ^2_{calc}	$\mathcal{P}(\chi^2 > \chi^2_{calc})$
$B \to X_s^L \gamma$	m_{ES}	15.153	0.713
	ΔE	14.669	0.743
$B \to X_s^H \gamma$	m_{ES}	9.664	0.961
	ΔE	17.7486	0.539
$B \to X_d^L \gamma$	m_{ES}	13.308	0.822
	ΔE	16.760	0.606
$B \to X_d^H \gamma$	m_{ES}	16.188	0.645
	ΔE	6.276	0.997

TABLE 4.17: χ^2 goodness of fit tests for fits to experimental data. Each test has 19 DOF and the probability of obtaining a χ^2 value greater than calculated given that many DOF is shown.

where n^k_{PDF} and n^k_{Data} are the respective projections for bin k and σ_{Data} is the positive (negative) Poisson error on the number of data events in that bin if $(n^k_{PDF} - n^k_{Data})$ is positive (negative). Given a calculated chi square value, χ^2_{calc}, and the number of degrees of freedom (DOF) in the calculation, which here is 19 for all cases, it is possible to calculate the probability that the χ^2 value would exceed χ^2_{calc} in an independent repetition of the test [63, 64]. This is to a good approximation the probability that the agreement between model and data will be worse if the analysis were repeated on an independent dataset. Table 4.17 lists the χ^2 value calculated for each of the eight data projections and associated probability. It can be seen the probabilities are all in excess of 0.5 suggesting that the models are in all cases good representations of the data.

4.6.4 Interpretation of Results

The above signal yields are converted into measured BFs through the relation,

$$\sum_{i=1}^{7} \mathcal{B}(B \to X_q^{i,m}\gamma) = \frac{n_{sig}}{2n_{B\overline{B}}\epsilon}, \tag{4.17}$$

where i is the mode index, $q = s, d$ identifies the decay flavour, $m = L, H$ identifies the mass region, n_{sig} is the signal yield from the fit, $n_{B\overline{B}}$ is the total number of $B\overline{B}$ pairs in data (470.9 million $B\overline{B}$ pairs) and ϵ is the signal event selection efficiency calculated from MC data. Signal efficiencies are calculated by taking the ratio of signal events surviving cuts in a given mass range to the total number of signal events generated with true hadronic mass in that mass range. Table 4.18 lists the signal yields, calculated efficiencies and estimated measured BFs with associated statistical errors.

Fit	Yield	MC Efficiency	Measured BF
$B \to X_s^L \gamma$	774±29	0.0447	$(1.84\pm0.07)\times10^{-5}$
$B \to X_s^H \gamma$	1015±39	0.0171	$(6.30\pm0.24)\times10^{-5}$
$B \to X_d^L \gamma$	33±9	0.0361	$(0.97\pm0.27)\times10^{-6}$
$B \to X_d^H \gamma$	59±15	0.0174	$(3.60\pm0.92)\times10^{-6}$

TABLE 4.18: Measured BF estimation and associated statistical error for each fit to experimental data.

Chapter 5

Post Fit Studies

The results from fits to experimental data are interpreted as BF measurements. This chapter describes the estimation of systematic uncertainties for each measured BF and then details the model-dependent extrapolation studies using MC data to estimate the inclusive BFs of both $B \to X_{s,d}\gamma$. Finally the ratio of these inclusive fractions is used to estimate a value of $|V_{td}/V_{ts}|$.

5.1 Experimental Systematic Uncertainties

Systematic uncertainties from the experimental techniques described in chapters 3 and 4 can be divided into five categories: Firstly differences in the reconstruction efficiency of primary particles within the *BABAR* framework between MC and experimental data; secondly differences in efficiencies between MC and experimental data of the cuts made during event reduction; thirdly uncertainties arising from fixing PDF parameters and event hypothesis yields in the MLF to experimental data; fourthly uncertainties in the models used to generate signal MC data, in particular the KN photon energy spectrum and JETSET phase space decay distribution of hadronic final states; finally the uncertainty in the estimate of the total number of $B\overline{B}$ pair events in experimental data. The error for this last category is given in section 3.1.1 and corresponds to a 0.6% uncertainty on each measured BF.

5.1.1 Reconstruction Efficiencies

Lists of reconstructed charged tracks and neutral clusters accessed from the central event store have different reconstruction efficiencies in MC and experimental data. This is because the GEANT4 simulation of the *BABAR* detector does not always accurately model

Fit Class	μ_{trk}	$\mu_{trk}\sigma_{trk}$
$B \to X_s^L \gamma$	1.43	0.30%
$B \to X_s^H \gamma$	2.31	0.49%
$B \to X_d^L \gamma$	1.57	0.33%
$B \to X_d^H \gamma$	1.93	0.41%

TABLE 5.1: Estimates of the charged track reconstruction systematic uncertainty for each experimental data fit.

the response of sub-detectors used in the reconstruction of those particle interactions. Dedicated groups within the *BABAR* collaboration study these differences with MC and experimental data control samples and advise analysts what uncertainty should be applied. The following summarises how the results of these dedicated *BABAR* studies have been applied to the reconstructed signal states in this analysis.

5.1.1.1 Charged Tracks

Differences between the reconstruction efficiencies of charged tracks in MC and experimental data are determined from a dedicated study carried out by the *BABAR* tracking group. The analysis is described in [68] and uses τ pair decays, where one τ decays leptonically and the subsequent charged lepton identifies the event. The recoiling τ must have an odd number of charged daughter particles. The study uses events where the recoiling τ has three charged daughters, two of which are charged pions either from a $\rho^0 \to \pi^+\pi^-$ decay or where both pions have total invariant mass greater than $300\,\mathrm{MeV}/c^2$ and opposite charge to the lepton from the other τ decay. These two pions and the charged lepton are identified in event reconstruction through strict PID, geometrical and kinematic criteria. The remaining charged daughter is either detected or not and the difference in its detection efficiency between MC and experimental data is assigned as the systematic uncertainty for charged track reconstruction. To account for the average multiplicity of these events being lower than that of a $B\overline{B}$ event a further correction factor is applied to the systematic uncertainty. For the datasets and reconstruction framework software used in this analysis the *BABAR* tracking group advise assigning an uncertainty of σ_{trk}=0.21% per track from the `GoodTracksLoose` list after the multiplicity correction is applied. For each of the fits described in chapter 4 the product of this uncertainty with the mean track multiplicity of correctly reconstructed MC signal decays after all event reduction cuts, μ_{trk}, is therefore applied as the total systematic uncertainty due to track reconstruction. These are summarised in table 5.1.

	$B \to X_s^L \gamma$	$B \to X_s^H \gamma$	$B \to X_d^L \gamma$	$B \to X_d^H \gamma$
μ_{π^0}	0.29	0.30	0.45	0.48
$\mu_{\pi^0}\sigma_{\pi^0}$	0.9%	0.9%	1.3%	1.4%
μ_η	-	0.04	-	0.03
$\mu_\eta\sigma_\eta$	-	0.1%	-	0.1%
σ_γ	0.7%	0.7%	0.7%	0.7%
σ_{neu}	1.6%	1.7%	2.0%	2.2%

TABLE 5.2: Estimates of the neutral particle reconstruction systematic uncertainty for each experimental data fit.

5.1.1.2 Neutral Particles

Studies performed by the *BABAR* neutral particles group are used to determine reconstruction efficiency differences in MC and experimental data for lists of photons as well as reconstructed π^0 and η mesons. The analysis described in [69] uses control samples of $\tau \to \pi\nu$ and $\tau \to \rho\nu$ decays to compare the reconstruction efficiencies of π^0 mesons and single photons. Consequently the *BABAR* neutral group recommend assigning a systematic uncertainty of σ_{π^0}=3.0% per π^0 from the `Pi0DefaultMass` list and 1.8% per single photon from the `GoodPhotonLoose` list, therefore giving an uncertainty per η meson of σ_η=3.6%, where $\eta \to \gamma\gamma$. The study described in [70] compares samples of $e^+e^- \to \mu^+\mu^-\gamma$ events in MC and experimental data to determine differences in reconstruction efficiency of high energy photons. Using the results of this study the *BABAR* neutral group advise assigning an uncertainty of σ_γ=0.7% for photons with a laboratory frame energy greater than 1 GeV.

Table 5.2 summarises the total systematic uncertainty, σ_{neu}, calculated for neutral particle reconstruction in each of the fits described in chapter 4. The π^0 and η uncertainties are calculated respectively as $\mu_{\pi^0}\sigma_{\pi^0}$ and $\mu_\eta\sigma_\eta$, where μ_M is the mean multiplicity of meson M in correctly reconstructed signal MC decays after event reduction cuts. As all neutral particle reconstruction uncertainties arise from photon detection in the EMC they are 100% correlated and therefore the total systematic uncertainty is a linear combination of those due to individual particles.

5.1.2 Event Reduction Cut Efficiencies

Imperfections in the detector simulation can lead to differences in distributions of variables between MC and experimental data. In particular any difference in a variable distribution on which a cut is applied during event reconstruction and subsequent reduction can lead to a systematic difference in the signal efficiency estimated from MC

data and the true cut efficiency in experimental data. The calculated BFs assume the MC efficiency accurately describes the data so an uncertainty must be applied for each variable which may negate this assumption. All variables identified as a potential source of uncertainty are listed below and their corresponding systematic uncertainty estimated.

5.1.2.1 Photon Cuts

The reduction cuts placed on the high energy photon of reconstructed candidates are the π^0/η candidate veto, variables relating to EMC cluster quality, specifically the second moment and lateral moment and the cut on distance to the nearest track-matched cluster. Previous radiative penguin analyses using *BABAR* data have made extensive studies of such cuts and estimated the relative uncertainties which should be applied for each. This analysis does not significantly modify the implementation of these cuts so none of these studies are repeated. Instead the systematic uncertainties calculated in previous analyses are applied.

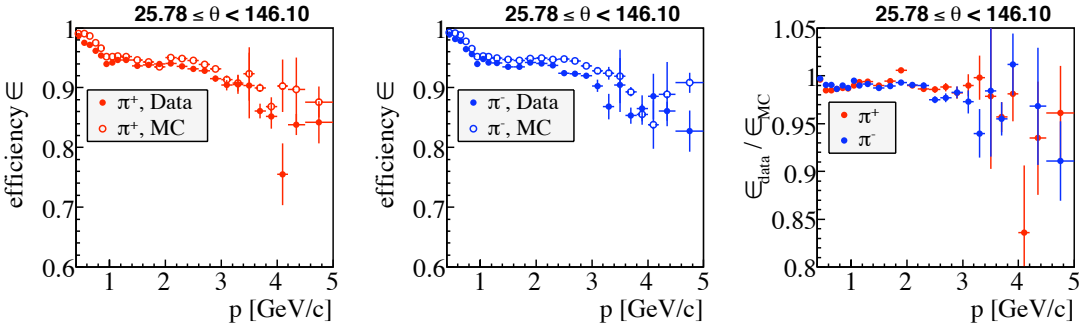
The analysis described in [71] investigates differences in performing the π^0/η candidate veto in MC and experimental data. A simulated signal photon with energy 2.5 GeV in the CM frame is embedded into different types of background MC and experimental data samples and the efficiency of the veto compared. The corresponding efficiencies are found to agree within 1% for each veto and hence this value is assigned as a systematic error. The errors are added linearly giving a total systematic uncertainty of 2% for these cuts. The same analysis also studies the effect of the track distance cut by embedding photon clusters extracted from $e^+e^- \to e^+e^-\gamma$ events into generic $B\overline{B}$ MC data and experimental data. An efficiency difference of 2% in applying the distance cut on these two samples is observed and applied as a systematic uncertainty.

The analysis described in [59] investigates photon cluster quality variables by comparing control samples of $e^+e^- \to \mu^+\mu^-\gamma$ decays in MC and experimental data. Only the second moment cut is found to show an efficiency difference, the magnitude of which is assigned as a systematic error of 2.1%. The cluster quality cut uncertainty is added in quadrature with the distance cut uncertainty to give an combined photon quality systematic error of 2.9%.

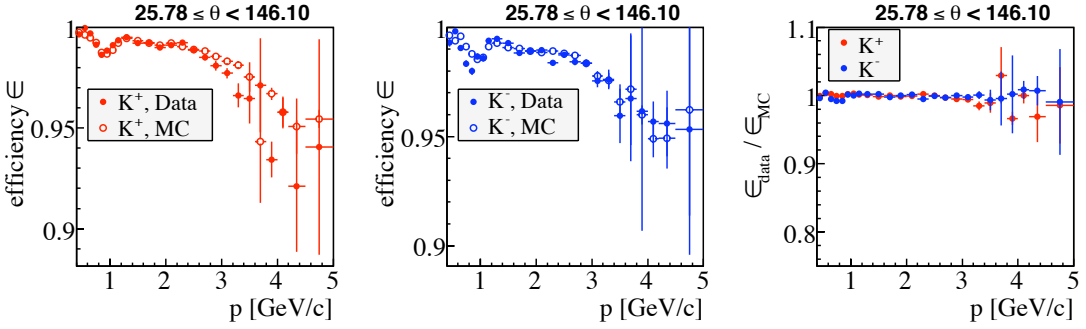
5.1.2.2 Charged Track PID

The *BABAR* PID group perform studies to identify differences in PID selector efficiencies between MC and experimental data as a function of track momentum. These studies use MC data events, from the control samples described in section 3.2.5, not used to

train the classifiers and their experimental data equivalents. Figure 5.1 shows the results of these efficiency studies for the PID selectors used to identify signal candidates during event reduction, `pionKMTight` and `kaonBDTNotAPion`. For this analysis signal candidates are generally reconstructed from charged tracks with momentum less than 2.5 GeV/c in the laboratory frame. The associated average systematic uncertainty for each selector are therefore taken from these plots to be 1.5% and 0.5% for `pionKMTight` and `kaonBDTNotAPion` respectively. Corrections for potential differences in event topology and track qualities between signal events and those used in the above control samples are not provided by the PID group; however, they advise that these are unlikely to be less than 1%. Consequently an additional uncertainty of 1% is added in quadrature for each selector giving an overall uncertainty per track of σ_π=1.8% and σ_K=1.1% for `pionKMTight` and `kaonBDTNotAPion` respectively.



(a) Efficiency comparisons for true pions passing the `pionKMTight` selector. From the plot of $\epsilon_{data}/\epsilon_{MC}$ an average systematic uncertainty of 1.5% per pion with momentum up to 2.5 GeV/c is assigned.



(b) Efficiency comparisons for true kaons passing the `kaonBDTNotAPion` selector. From the plot of $\epsilon_{data}/\epsilon_{MC}$ an average systematic uncertainty of 0.5% per kaon with momentum up to 2.5 GeV/c is assigned.

FIGURE 5.1: Plots from the *BABAR* PID group showing efficiency comparisons of PID selector performance between MC and experimental data control samples for (a) `pionKMTight` and (b) `kaonBDTNotAPion`.

Signal candidates reconstructed in this analysis have PID selector requirements imposed on all charged tracks used in their reconstruction. Kaon candidates must pass the `kaonBDTNotAPion` selector and fail the `pionKMLoose` selector. Pion candidates must pass the `pionKMTight` selector and fail the `kaonBDTNotAPion` selector. Charged tracks

	$B \to X_s^L \gamma$	$B \to X_s^H \gamma$	$B \to X_d^L \gamma$	$B \to X_d^H \gamma$
μ_π	0.71	1.31	1.57	1.93
$\mu_\pi \sigma_\pi$	1.3%	2.4%	2.8%	3.5%
μ_K	1.00	1.00	-	-
$\mu_K \sigma_K$	1.1%	1.1%	-	-
σ_{PID}	2.4%	3.5%	2.8%	3.5%

TABLE 5.3: Estimates of the PID selector systematic uncertainty for each experimental data fit.

passing the `pionKMTight` selector are a subset of those passing the `pionKMLoose` selector. One track in 10^6 from the `GoodTracksLoose` list is found to pass both the `pionKMTight` and `kaonBDTNotAPion` in both signal MC and experimental data. For the purpose of applying a PID systematic uncertainty it is therefore assumed that the selectors are essentially mutually exclusive and that any track which fails one of these selectors due to PID inefficiency will not simultaneously be accepted by the other selector. Consequently the pion selector uncertainty need only be applied to tracks identified as pions and the kaon selector uncertainty only to tracks identified as kaons. There is significant overlap in many of the input variables used in each selector so a 100% correlation between the pion and kaon selector inefficiencies are assumed. The total PID systematic uncertainty for a class of signal events, σ_{PID}, is therefore taken to be,

$$\sigma_{PID} = \sigma_\pi \mu_\pi + \sigma_K \mu_K, \tag{5.1}$$

where μ_π and μ_K are the mean number of respective charged pions and kaons in signal MC passing all event reduction cuts. Table 5.3 summarises the PID systematic uncertainties assigned to each of the signal event classes.

5.1.2.3 Hadronic Candidate and Event Topology Cuts

Event reduction cuts placed on variables relating to the hadronic candidate or event topology are considered individually for potential systematic differences in the signal selection efficiency between MC and experimental data. Four variables are identified as potential sources of uncertainty due to non-uniform distributions in MC data of correctly reconstructed candidates. These are the minimum daughter momentum cut, the vertex χ^2 probability cut, the decay length cut which forms part of the K_S^0 veto in $B \to X_d \gamma$ candidate reconstruction and the NN cut.

Systematic uncertainties from placing cuts on these variables are determined by taking advantage of the dominant signal event contributions in fits to both the high and low

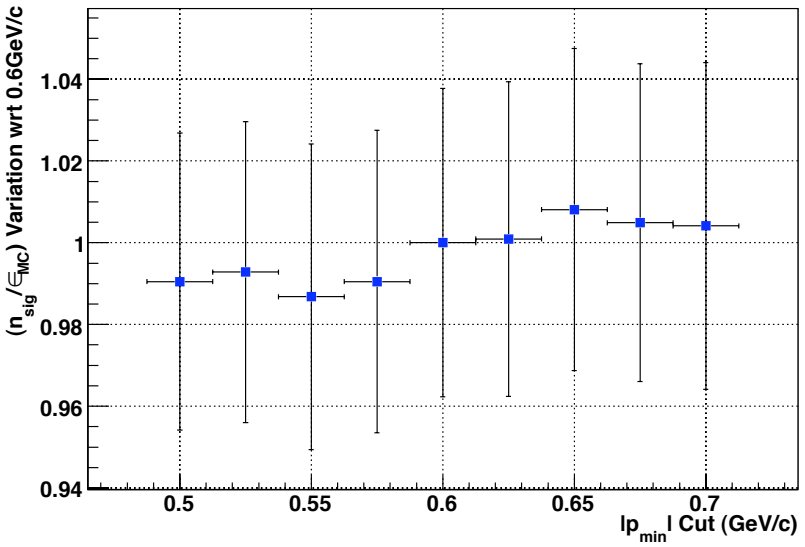
mass region $B \to X_s\gamma$ candidate experimental data. In all cases the $B \to X_s\gamma$ candidate distributions for these variables are assumed to be a good representation of the corresponding $B \to X_d\gamma$ candidate distributions. Therefore the systematic uncertainty estimated from $B \to X_s\gamma$ data is also directly applicable to $B \to X_d\gamma$ data where the signal contribution in experimental data is not significant enough to allow an independent investigation. For each variable the cut is varied locally in MC and experimental data. The PDF parameterisations and fits to experimental data are then repeated for the new cut value and the corresponding value of n_{sig}/ϵ_{MC} calculated. This is then normalised to the value of n_{sig}/ϵ_{MC} from the default cut value and the average deviation relative to the default value is then assigned as the systematic error for that variable.

Figure 5.2 shows the variation in n_{sig}/ϵ_{MC} with respect to the default cut value for high and low mass $B \to X_s\gamma$ candidate data for the minimum daughter momentum cut, $|p_{min}|$. The cuts are varied by $\pm 0.1\,\mathrm{GeV}/c$ in steps of $0.025\,\mathrm{GeV}/c$, which corresponds to the step size used in cut optimisation. Across this range an average local deviation of 1.0% and 1.7% are observed for $B \to X_s^L\gamma$ and $B \to X_s^H\gamma$ candidate data respectively.

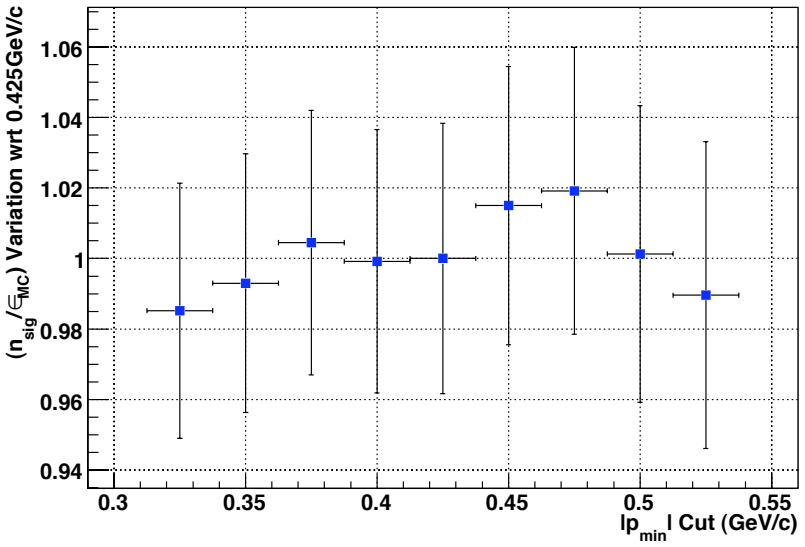
Figure 5.3 shows the variation in n_{sig}/ϵ_{MC} with respect to the default cut value for high and low mass $B \to X_s\gamma$ candidate data for the vertex χ^2 probability cut. The cuts are varied from 0.1% to 2.5% and across this range an average local deviation of 0.5% and 1.0% are observed for $B \to X_s^L\gamma$ and $B \to X_s^H\gamma$ candidate data respectively.

Figure 5.4 shows the variation in n_{sig}/ϵ_{MC} with respect to the default cut value for high and low mass $B \to X_s\gamma$ candidate data for the vertex decay length cut, x_{DEC}. The cuts are varied by $\pm 0.1\,\mathrm{cm}$ and across this range an average local deviation of 0.2% and 0.4% are observed for $B \to X_s^L\gamma$ and $B \to X_s^H\gamma$ candidate data respectively. This systematic uncertainty only applies to $B \to X_d\gamma$ candidate data as the decay length cut forms part of the K_S^0 veto which is only applied to these data.

Figure 5.5 shows the variation in n_{sig}/ϵ_{MC} with respect to the default cut value for high and low mass $B \to X_s\gamma$ candidate data for the NN cut. The cuts are nominally varied by ± 0.1; however, the default cut value of 0.92 for the low mass NN cut restricts the upper limit of variation to 0.98. The cuts are varied in steps of 0.02 which corresponds to the step size used for this variable during the cut optimisation. Across this range an average local deviations of 2.5% and 3.1% are observed for $B \to X_s^L\gamma$ and $B \to X_s^H\gamma$ candidate data respectively.

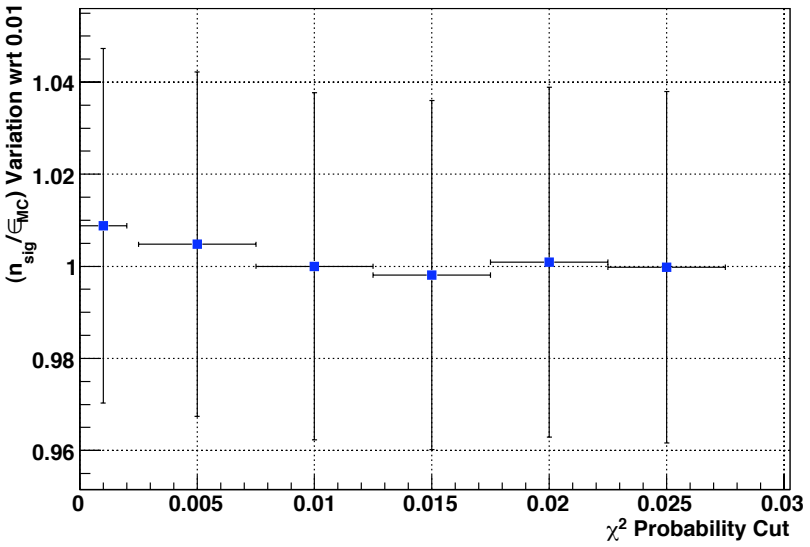


(a) $B \to X_s^L\gamma$ data, the default cut is $0.6\,\mathrm{GeV}/c^2$. From these variations an average systematic uncertainty of 1.0% is assigned.

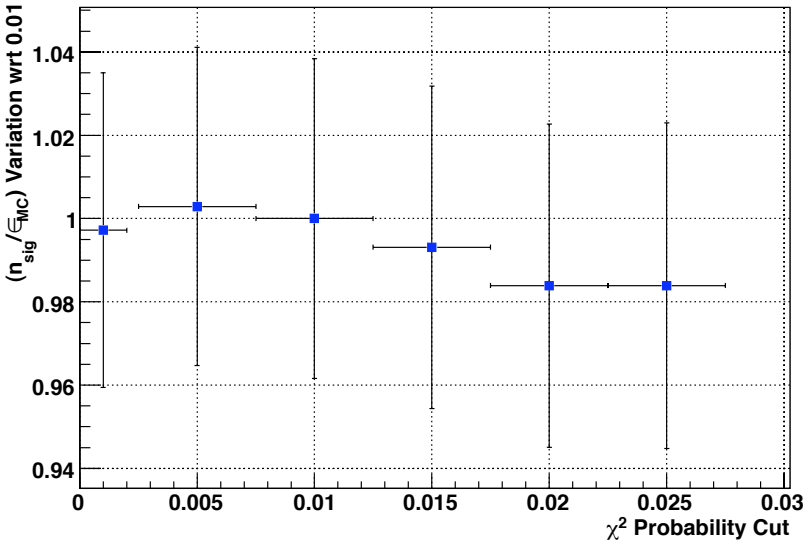


(b) $B \to X_s^H\gamma$ data, the default cut is $0.425\,\mathrm{GeV}/c^2$. From these variations an average systematic uncertainty of 1.7% is assigned.

FIGURE 5.2: Systematic variation of n_{sig}/ϵ_{MC} as a function of $|p_{min}|$ cut, with respect to the default analysis cut for fits to (a) $B \to X_s^L\gamma$ data (b) $B \to X_s^H\gamma$ data. The vertical error bars represent the statistical uncertainty in fits to experimental data of the parameter n_{sig}.

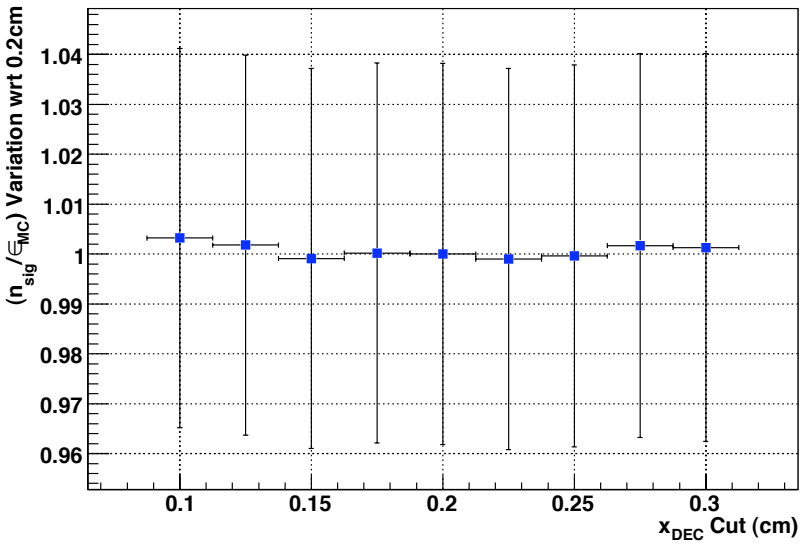


(a) $B \to X_s^L\gamma$ data, the default cut is 0.01. From these variations an average systematic uncertainty of 0.5% is assigned.

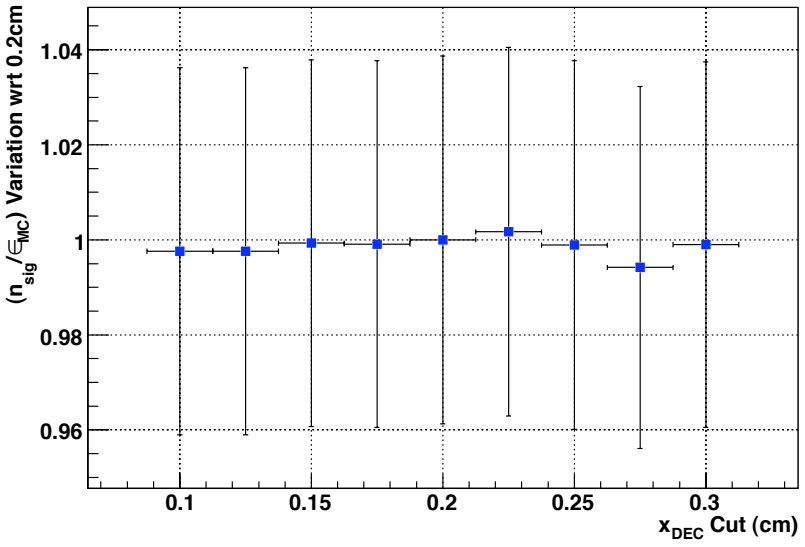


(b) $B \to X_s^H\gamma$ data, the default cut is 0.01. From these variations an average systematic uncertainty of 1.0% is assigned.

FIGURE 5.3: Systematic variation of n_{sig}/ϵ_{MC} as a function of χ^2 probability cut, with respect to the default analysis cut for fits to (a) $B \to X_s^L\gamma$ data (b) $B \to X_s^H\gamma$ data. The vertical error bars represent the statistical uncertainty in fits to experimental data of the parameter n_{sig}.

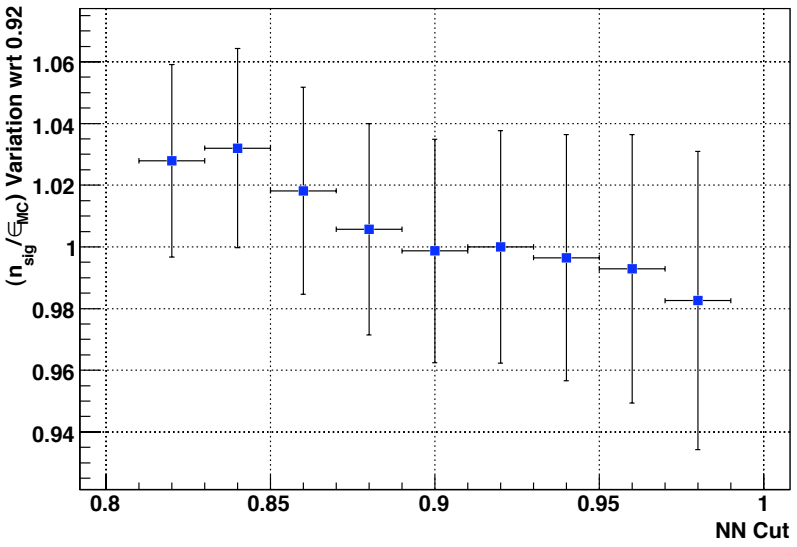


(a) $B \to X_s^L \gamma$ data, the default cut in $B \to X_d^L \gamma$ data is 0.2 cm. From these variations an average systematic uncertainty of 0.2% is assigned.

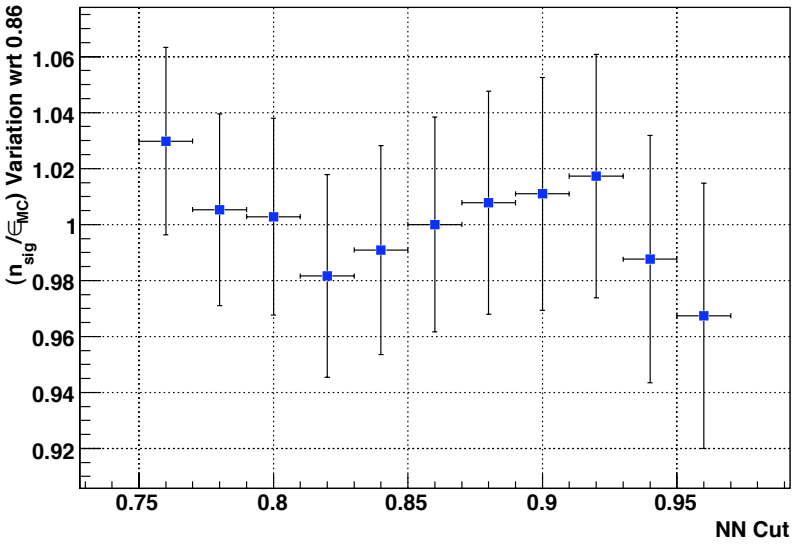


(b) $B \to X_s^H \gamma$ data, the default cut in $B \to X_d^H \gamma$ data is 0.2 cm. From these variations an average systematic uncertainty of 0.4% is assigned.

FIGURE 5.4: Systematic variation of n_{sig}/ϵ_{MC} as a function of x_{DEC} cut, with respect to the default analysis cut for fits to (a) $B \to X_s^L \gamma$ data (b) $B \to X_s^H \gamma$ data. The vertical error bars represent the statistical uncertainty in fits to experimental data of the parameter n_{sig}.



(a) $B \to X_s^L\gamma$ data, the default cut is 0.92. From these variations an average systematic uncertainty of 2.5% is assigned.



(b) $B \to X_s^H\gamma$ data, the default cut is 0.86. From these variations an average systematic uncertainty of 3.1% is assigned.

FIGURE 5.5: Systematic variation of n_{sig}/ϵ_{MC} as a function of NN cut, with respect to the default analysis cut for fits to (a) $B \to X_s^L\gamma$ data (b) $B \to X_s^H\gamma$ data. The vertical error bars represent the statistical uncertainty in fits to experimental data of the parameter n_{sig}.

5.1.3 Fit Uncertainties

Event hypothesis yields and PDF parameters which are not floated in the EMLF to experimental data can lead to systematic differences in MC and experimental data. These difference can arise if the simulation does not accurately reflect the fit variable distribution shapes and normalisations or if the size of the MC data sample used to determine fixed parameters leads to significant statistical uncertainties on those estimated parameter values. Furthermore biases observed in toy MC studies indicate potential systematic uncertainties on the signal yield due to variables floated in the combined fit to data. Uncertainties must therefore be estimated for all potential sources of bias and for variables held constant in the fit, the details of which are described below.

5.1.3.1 Fixed PDF Parameters

Fixing PDF parameters assumes that the PDF distributions in MC data are a good model of the corresponding experimental data. However, the limited sample sizes of each MC event class results in an associated statistical error for each fitted parameter. These are the errors shown in tables 4.2, 4.3, 4.4 and 4.5. Consequently all fits to data are repeated for each fixed parameter, varying its value to the extremes of corresponding statistical error. The average variation in n_{sig} is then assigned as the systematic uncertainty due to that parameter. The systematic errors for all fixed variables in a given PDF are then combined, taking any correlations between parameters in the PDF fit into account. The correlation matrix between fixed PDF parameters, $\mathcal{C}$, is calculated from the fits to MC data discussed in section 4.4. The average systematic variations on n_{sig} are expressed as an error vector $\boldsymbol{\epsilon}$ whose size is equal to the number of fixed PDF parameters. This results in an overall systematic uncertainty for that PDF, ϵ', given by,

$$\epsilon' = \sqrt{\boldsymbol{\epsilon}^T \mathcal{C} \boldsymbol{\epsilon}}. \tag{5.2}$$

Systematic uncertainties for all PDFs in a given fit to data are combined in quadrature to give an overall uncertainty due to all fixed parameters in that fit.

Tables 5.4 and 5.5 show the calculated uncertainties for each fixed parameter from m'_{ES} and ΔE PDF distributions respectively, as well as the associated combined uncertainty for those variables. This gives a total systematic uncertainty of 0.3% for $B \to X_s^L \gamma$ data, 2.9% for $B \to X_s^H \gamma$ data, 5.9% for $B \to X_d^L \gamma$ data and 3.6% for $B \to X_d^H \gamma$ data due to fixed PDF parameters.

The 2D histogram PDFs which parameterise cross feed distributions do not have parameters to vary. However, each bin of the PDF does have an associated statistical error

Hypothesis	Parameter	$B \to X_s^L \gamma$	$B \to X_s^H \gamma$	$B \to X_d^L \gamma$	$B \to X_d^H \gamma$
Signal	μ	0.01%	0.01%	0.03%	0.23%
	σ	0.02%	0.11%	0.10%	0.48%
	α	0.04%	0.20%	0.14%	0.70%
	n	0.03%	0.12%	0.08%	0.32%
	combined	0.02%	0.10%	0.09%	0.47%
X_s PID cross-feed	μ	-	-	0.12%	0.03%
	σ	-	-	0.40%	0.20%
	α	-	-	0.62%	0.28%
	n	-	-	0.41%	0.16%
	combined	-	-	0.36%	0.13%
Generic $B\overline{B}$	ξ	0.25%	1.17%	2.57%	0.34%
	p	0.22%	2.55%	2.74%	0.14%
	μ	0.04%	0.11%	0.71%	0.88%
	σ	0.12%	1.03%	1.65%	1.81%
	f	0.25%	2.91%	3.69%	3.60%
	combined	0.14%	0.74%	2.24%	3.04%
Continuum	ξ	0.68%	-	12.49%	-
	p	0.75%	2.77%	14.26%	0.16%
	combined	0.26%	2.77%	5.02%	0.16%
	Total Uncertainty	0.30%	2.87%	5.51%	3.08%

TABLE 5.4: Systematic uncertainties from fixing m'_{ES} PDF parameters in fits to experimental data.

so in principle each bin could be varied by the range of this error and the fit repeated. The combination in quadrature of the average variation in n_{sig} from all bins would then give an overall uncertainty for the PDF. The fit framework used in this analysis does not allow such variations to be made without significant modification, therefore this uncertainty is not evaluated. The cross feed projection plots for fitted MC data in chapter 4 show that in the signal region these variations are generally small, therefore this is not expected to be a dominant systematic uncertainty. Neglecting this error is consistent with the previous version of this analysis [1].

5.1.3.2 Agreement Between MC and Experimental Data

The above systematic variations only account for uncertainties in fixed PDF parameters due to statistical limits of the MC data used for their determination. In the limit of infinite MC data points such uncertainties would be zero. An additional source of systematic uncertainty can arise if a parameter value is fundamentally different in the

Hypothesis	Parameter	$B \to X_s^L \gamma$	$B \to X_s^H \gamma$	$B \to X_d^L \gamma$	$B \to X_d^H \gamma$
Signal	μ	0.01%	0.04%	0.41%	0.03%
	σ_L	-	-	1.14%	0.83%
	σ_R	-	-	0.28%	0.90%
	α_L	0.03%	0.11%	0.19%	0.54%
	α_R	0.01%	0.06%	0.07%	0.44%
	combined	0.03%	0.11%	0.88%	0.59%
X_s PID cross-feed	μ	-	-	1.55%	0.75%
	σ_L	-	-	1.28%	0.56%
	σ_R	-	-	0.74%	0.42%
	α_L	-	-	0.61%	0.28%
	α_R	-	-	0.02%	0.05%
	combined	-	-	0.52%	0.24%
Generic $B\overline{B}$	a_1	0.01%	0.03%	0.08%	0.20%
	a_2	0.08%	0.17%	1.69%	1.82%
	combined	0.07%	0.15%	1.74%	1.69%
Continuum	a_1	0.01%	-	0.04%	-
	combined	0.01%	-	0.04%	-
	Total Uncertainty	0.08%	0.19%	2.02%	1.81%

TABLE 5.5: Systematic uncertainties from fixing ΔE PDF parameters in fits to experimental data.

MC data and experimental data due to modelling errors. However, such an uncertainty will only be significant if the difference is much greater than the statistical variations already considered.

Signal PDF distributions use the $B \to X_s\gamma$ fits to data as a control sample to compare MC and experimental data distributions. In these data the contribution of signal events dominates. Parameters which are known to not be modelled well in the MC data, specifically the widths of ΔE distributions, are measured and used to correct the signal shapes in $B \to X_d\gamma$ fits to data. The remaining fixed PDF parameters for these distributions show no significant difference between MC and experimental data beyond the statistical uncertainty of those parameters in the control sample.

Cross-feed distributions modelled as 2D histograms do not depend on fixed PDF parameters; however, the shape of these histograms could be expected to show systematic differences between MC and experimental data particularly for decays generated from JETSET. The application of weights to the $B \to X_s\gamma$ MC data discussed in section 4.2 can indicate the potential for such differences. In all cases the difference in shape between weighted and unweighted bins for high mass $B \to X_s\gamma$ MC data histograms

were found to be negligible. The only significant differences observed were in the overall normalisations which are discussed below.

Variations of fixed PDF parameters for generic B background PDFs contribute significantly to the overall fit uncertainty. Tables 5.4 and 5.5 show that this can be attributed to the peaking component in m'_{ES} as generally the Argus function parameter variations in m'_{ES} for this PDF are anti-correlated and cancel. The peaking backgrounds typically from events with low multiplicity B decays which have a similar final state to the signal mode reconstructed. The previous version of this analysis [1] carried out dedicated studies using charmless $B \to X\pi^0$ and $B \to X\eta$ MC samples to model such backgrounds with independent PDF distributions. No significant difference was observed between parameters of the peaking components from these distributions with those peaking backgrounds in other generic B decays. Therefore in this analysis they were modelled with a combined PDF distribution determined from the generic sample without such events vetoed. Despite a larger generated generic sample in this analysis, the optimised cuts are generally tighter than before leading to similar numbers of events used to fit the PDFs for this background. The resulting statistical variations on fixed parameters are hence comparable. The previous analysis determined that these variations were in excess of any underlying systematic differences between the MC and experimental distribution parameters and this assumption is therefore made here too. Indeed from the statistical variations only two parameters from the m'_{ES} distributions are of specific concern due to the sensitivity of n_{sig} to their respective values: the Gaussian width and relative fraction of Argus normalisation to Gaussian normalisation. The fraction parameter uncertainty is to an extent addressed in the normalisation variation discussed below. This is because the change in Gaussian normalisation will generally be absorbed by the similarly peaking signal distribution, hence biasing n_{sig}, whereas the change in Argus normalisation will generally be absorbed by the floated continuum Argus. Indeed variations in the Generic B normalisations are anti-correlated with n_{sig}. Knowledge of potential uncertainties from the Gaussian width can be determined from signal distributions. The signal width in m'_{ES} is well modelled by the MC data as shown from the fits to $B \to X_s\gamma$ experimental data. This width represents the detector resolution limit of this variable for correctly reconstructed candidates. The possibility of a peaking background having width narrower than this limit is physically unreasonable. With the exception of high mass $B \to X_s\gamma$ candidates the negative statistical variation of the fixed peaking background width takes it below the signal width as can be seen from tables 4.2 and 4.4. Hence the variations made are already likely to be too conservative; in fact the variations in n_{sig} from varying the the Gaussian widths are generally asymmetric with the negative, non-physical extreme giving the larger uncertainty. This is because a wider peaking component leads to this background having a

Yield	$B \to X_s^L \gamma$	$B \to X_s^H \gamma$	$B \to X_d^L \gamma$	$B \to X_d^H \gamma$
n_{scf}	0.96%	3.39%	1.12%	1.92%
n_{mcf}	1.99%	0.86%	0.79%	0.52%
n_{PID}^{Xs}	-	-	3.11%	1.04%
n_{scf}^{Xs}	-	-	1.61%	3.21%
n_{mcf}^{Xs}	-	-	2.97%	0.24%
n_{GenB}	0.61%	2.78%	7.33%	8.23%
Total Uncertainty	2.29%	4.47%	8.76%	9.12%

TABLE 5.6: Systematic uncertainties from fixing event hypothesis yields in fits to experimental data.

more continuum-like shape, thus biasing the signal yield less.

The continuum PDF fixed parameter variations are only a significant concern in the low mass $B \to X_d\gamma$ fit to experimental data where it was not possible to float the Argus slope parameter due to toy study fits failing when this freedom was allowed in the combined yield fit. However, the statistical variations made are relatively large due to low MC data sample sizes after event selection cuts. The statistical variations are therefore assumed to be sufficient to parameterise potential systematic uncertainties.

5.1.3.3 Fixed Event Class Normalisations

In fits to experimental data the relative normalisations of cross feed and generic B PDF contributions are fixed, thereby assuming the MC data accurately represent the relative reconstruction efficiencies of candidates within these event classes. However, not all of the contributions are necessarily well modelled, hence uncertainties exist in these normalisations. To estimate the associated systematic uncertainty the fits to data are repeated assuming a predetermined 'conservative' variation in each normalisation and taking the average variation in n_{sig} as the uncertainty for that event class. Individual uncertainties from each event class are combined in quadrature to give an overall systematic uncertainty for a particular data fit.

Table 5.6 summarises the uncertainties estimated from variations made to each event class normalisation in fits to experimental data. With the exception of PID cross feed, cross feed contributions are varied by ±30%. This is because the application of weights to $B \to X_s\gamma$ MC data discussed in section 4.2 do not vary the relative normalisations by more than this amount. This is therefore used as an estimate of how much normalisations could vary for all cross feed distributions. The iterative updating of these yields in fits to data means that this is likely to be conservative variation as in those iterations all cross-feed normalisations converged. The PID cross feed only contributes to $B \to X_d\gamma$

Parameter	$B \to X_s^L \gamma$	$B \to X_s^H \gamma$	$B \to X_d^L \gamma$	$B \to X_d^H \gamma$
n_{sig}	0.5%	0.1%	2.9%	1.0%
n_{cont}	0.0%	0.1%	0.1%	0.2%
σ_L	0.2%	0.8%	-	-
σ_R	0.0%	0.2%	-	-
ξ	-	0.0%	-	0.7%
a_1	-	0.0%	-	0.0%
Combined Uncertainty	0.6%	0.9%	2.9%	1.3%

TABLE 5.7: Systematic uncertainties from bias on parameters floated in fits to experimental data.

data samples and is formed of correctly reconstructed modes which have been measured in the corresponding $B \to X_s\gamma$ data to a statistical accuracy of lower than 5% in both the high and low mass regions. Control sample studies by the PID group show that the rate of charged kaons being misidentified as pions is less than 10% for all momenta with reasonable agreement between MC and experimental data. For these reasons the PID cross feed normalisation is varied by $\pm 10\%$ or ± 1 event, whichever is greater.

Studies of generic B MC data find the generated events of such candidates are generally not dominated by a particular B decay. As discussed above, this background was parameterised differently in the previous version of this analysis. The normalisation is varied by $\pm 50\%$. This is to conservatively allow for the possibility of a significantly larger or smaller peaking m'_{ES} background component which, as seen from the statistical variations above, is potentially a very significant systematic uncertainty. It should be noted that it is potentially possible to make a more accurate estimate of the agreement of generic B background normalisation between MC and experimental data. The hadronic mass distribution of $B \to X_s^H \gamma$ candidates in generic B events contains a peak at the D^0 where the D^0 decay is identical to the an X_s final state, has been correctly reconstructed but combined with a photon from elsewhere in the event. The corresponding peak is also observed in the hadronic mass distribution of the experimental data. A detailed long term study may therefore be possible using this peak to better understand any difference between MC and experimental data normalisation.

5.1.3.4 Fit Bias

Bias on the signal yield from parameters floated in fits to data are a potential source of systematic uncertainty. The embedded toy studies pulls calculated in section 4.5.2 are used to correct floated parameters for biases in their own value in fits to experimental data. To evaluate the effect of such biases on signal yields each fit is repeated fixing a

given parameter to its bias corrected value and allowing all other parameters to float as normal. The resulting variation in the signal yield is taken as the uncertainty due to the bias from that parameter. The uncertainty on signal yield itself is simply the relative size of its own bias correction. Table 5.7 lists the uncertainties calculated from this method for each fit to data. The combined uncertainty takes correlations between the floated variables into account by using correlation matrices calculated from the default fits to experimental data and applying (5.2).

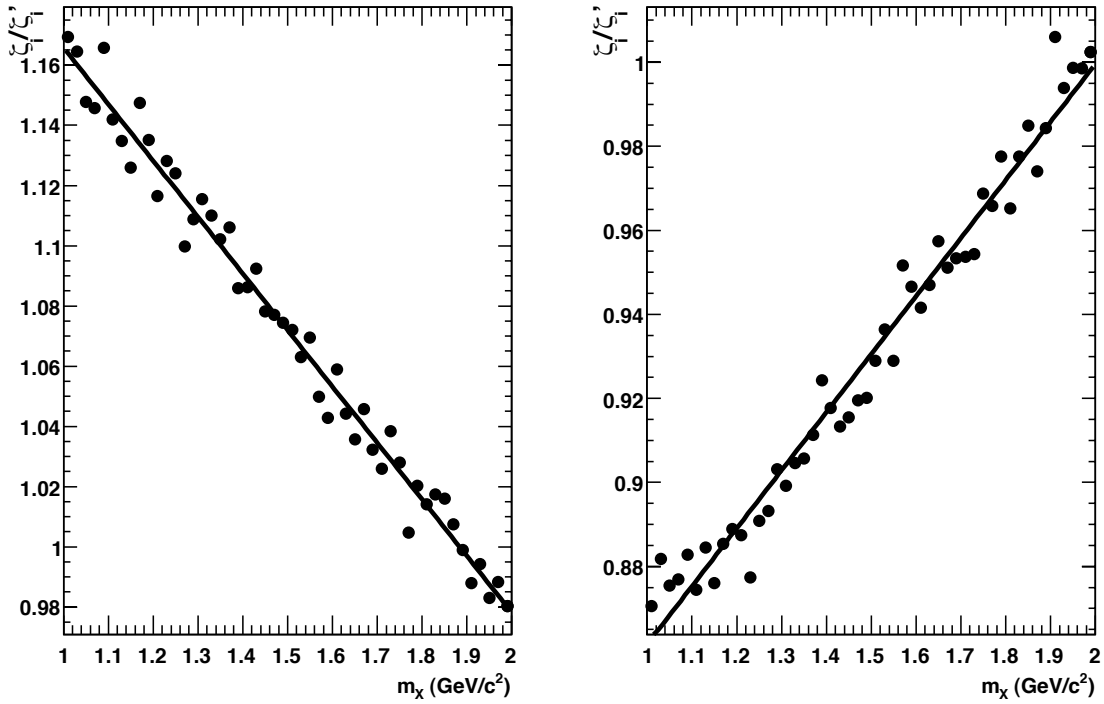
5.1.4 Signal Model Uncertainties

Assumptions made in modelling signal decay distributions in MC data can have associated systematic uncertainties if the model is not an accurate representation of true signal events or has any associated uncertainty itself. This section considers potential sources of error on the signal efficiency calculated from signal MC samples and quantifies the corresponding systematic uncertainties.

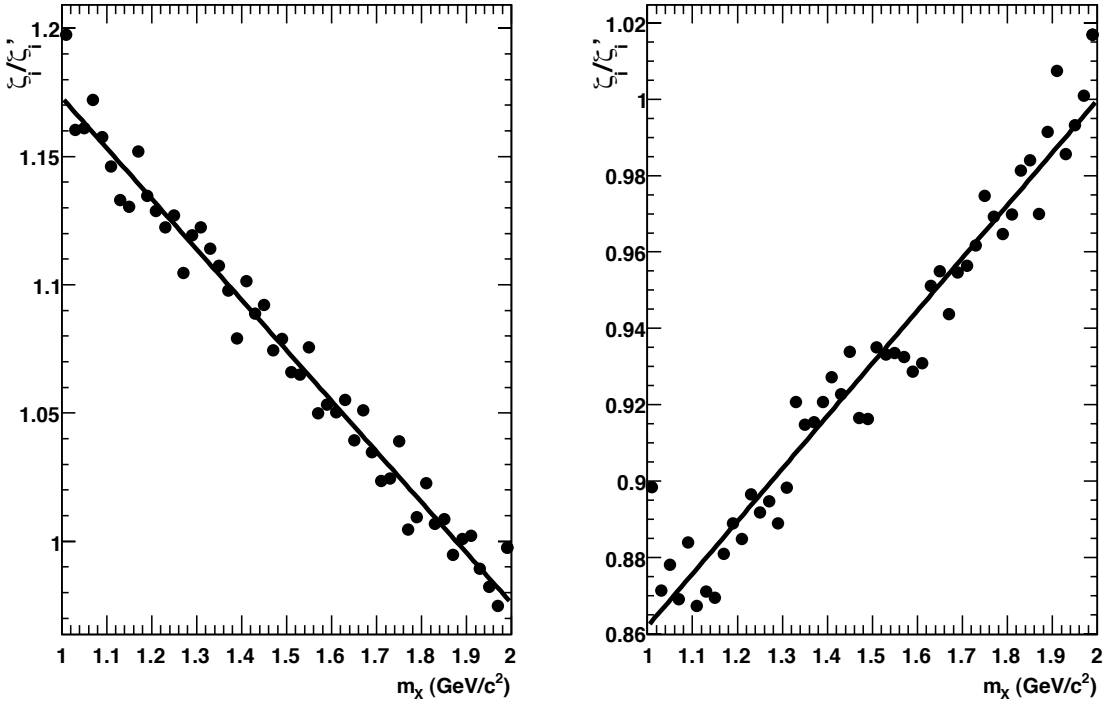
5.1.4.1 Photon Spectrum Model

Systematic variations to the KN photon spectrum parameterisation, given numerically in chapter 1, are used to estimate the associated uncertainty of the signal efficiency due to this model. Signal MC data with a total 2,122,000 events are generated for both $B \to X_s\gamma$ and $B \to X_d\gamma$ decays under each set of alternative KN parameters. From these data the relative contribution of signal states in the range $1.0 \leq m_X \leq 2.0\,\mathrm{GeV}/c^2$ before cuts are determined. To estimate the number of events passing all cuts under an alternative model the analysis is not repeated on these new data. Instead all generated MC events in the signal mass range, N_{sig}, are separated into 50 equal width bins of generated hadronic mass, each containing n_i (i=1..50) events. These are then normalised to the total number of events in the signal mass range, giving a proportion of events in each bin, $\zeta_i = n_i/N_{sig}$. The proportion of events in that bin are then compared to the corresponding proportion under the default KN parameterisation, ζ'_i, through the relation ζ_i/ζ'_i. This ratio of proportions is interpreted as a weight which, when applied to the default KN model, maps the default parameterisation onto the new parameterisation. Thus a weighted sum based on the generated hadronic mass of correctly reconstructed events passing all cuts in the default model gives the expected number of events which will pass all cuts under that alternative parameterisation.

Figure 5.6 shows plots of ζ_i/ζ'_i across the mass range for each alternative KN model for both $B \to X_s\gamma$ and $B \to X_d\gamma$ transitions. To simplify the calculation of weighted event sums, the weights are parameterised as a linear function in hadronic mass by



(a) $B \to X_s\gamma$ alternative KN mass weights and fitted linear polynomials. The left plot shows the weight distribution needed when using the default parameterisation, (m_b, μ_π^2)=(4.65 GeV/c^2,0.52 GeV^2), to emulate the parameterisation (4.70 GeV/c^2,0.45 GeV^2) and the right plot the corresponding distribution to emulate the parameterisation (4.60 GeV/c^2,0.60 GeV^2).



(b) $B \to X_d\gamma$ alternative KN mass weights and fitted linear polynomials. The left plot shows the weight distribution needed when using the default parameterisation, (m_b, μ_π^2)=(4.65 GeV/c^2,0.52 GeV^2), to emulate the parameterisation (4.70 GeV/c^2,0.45 GeV^2) and the right plot the corresponding distribution to emulate the parameterisation (4.60 GeV/c^2,0.60 GeV^2).

FIGURE 5.6: Calculated hadronic mass event weight distributions for emulation of alternative KN model parameterisations from the default in (a) $B \to X_s\gamma$ MC data and (b) $B \to X_d\gamma$ MC data. Each distribution is fitted with a linear polynomial function.

KN parameterisation	Weight Function	ϵ_{MC}
(4.65 GeV/c^2,0.52 GeV^2)	1.000	1.71%
(4.60 GeV/c^2,0.60 GeV^2)	0.733+0.148m_X	1.82%
(4.70 GeV/c^2,0.45 GeV^2)	1.272−0.176m_X	1.62%

TABLE 5.8: Fitted weight functions and calculated MC data efficiency for alternative KN parameterisations in $B \to X_s\gamma$ MC data.

KN parameterisation	Weight Function	ϵ_{MC}
(4.65 GeV/c^2,0.52 GeV^2)	1.000	1.74%
(4.60 GeV/c^2,0.60 GeV^2)	0.733+0.147m_X	1.85%
(4.70 GeV/c^2,0.45 GeV^2)	1.285−0.185m_X	1.64%

TABLE 5.9: Fitted weight functions and calculated MC data efficiency for alternative KN parameterisations in $B \to X_d\gamma$ MC data.

fitting 1D polynomials to the corresponding histograms. Tables 5.8 and 5.9 list the respective event weight functions and calculated MC data efficiencies for different KN models in $B \to X_s\gamma$ and $B \to X_d\gamma$ MC data. From these efficiency variations an average uncertainty of 5.8% and 6.0% are assigned to the measured $B \to X_s^H\gamma$ and $B \to X_d^H\gamma$ BFs respectively.

5.1.4.2 JETSET Phase Space Model

Uncertainties in the signal efficiency due to differences between the relative distribution of final states for the JETSET decay model in KN signal MC data and their true distribution in experimental data are evaluated for $B \to X_s^H\gamma$ and $B \to X_d^H\gamma$ decays.

As discussed in section 4.2, correction weights have been applied to $B \to X_s^H\gamma$ MC data before calculating the signal efficiency. Therefore the only remaining uncertainty in the distribution of these decays in experimental data arises from the statistical uncertainty of these measured weights. The weights and their associated uncertainties are given in table B.6 of appendix B. For each signal mode the applied weight is varied to the extremes of its quoted error while keeping the other signal weights constant. The total signal efficiency, ϵ_{MC}, is then recalculated and the average relative variation in the efficiency assigned as the associated uncertainty for that mode. The uncertainties for each mode are then added in quadrature to give an overall systematic uncertainty due to the relative distribution of signal states in $B \to X_s^H\gamma$ data. Table 5.10 list the relative uncertainty of each signal mode due to its weight. The combined uncertainty for all signal modes is 3.5%.

Signal Mode	Uncertainty
$B^0 \to K^+\pi^-\gamma$	1.0%
$B^0 \to K^+\pi^0\gamma$	0.5%
$B^0 \to K^+\pi^-\pi^+\gamma$	0.6%
$B^0 \to K^+\pi^-\pi^0\gamma$	1.2%
$B^0 \to K^+\pi^-\pi^+\pi^-\gamma$	0.9%
$B^0 \to K^+\pi^-\pi^+\pi^0\gamma$	2.7%
$B^0 \to K^+\eta\gamma$	0.7%
All Signal	3.5%

TABLE 5.10: Systematic uncertainties due to uncertainties in the weights applied to $B \to X_s^H\gamma$ signal MC data modes.

Calculating the corresponding uncertainty for $B \to X_d^H\gamma$ data is less straightforward as no information from experimental data exist and the default distribution of final states from JETSET decays is assumed to be correct when calculating ϵ_{MC}. Instead an alternative model of $B \to X_d^H\gamma$ final state distributions is proposed using the known weights from $B \to X_s^H\gamma$ data , applying them to the $B \to X_d^H\gamma$ data and then taking the resulting change in MC efficiency as an estimate of the size of systematic variation which could occur due to the difference in final state distributions in MC and experimental data.

The alternative model of signal mode contributions for $B \to X_d^H\gamma$ data is designed to reflect the relative contribution of different final states previously measured in $B \to X_s^H\gamma$ data. For these data the s quark produced in the radiative penguin transition is subsequently found in either a charged or neutral K meson. In substituting the s quark for a d quark the confined quark state substitutions listed in table 5.11 are assumed. From isospin symmetry, the $|d\bar{d}\rangle$ state contributes to the π^0, η and η' states in the pseudoscalar meson nonet. Under this scheme the meson states are given as [5],

$$\pi^0 = \frac{(|u\bar{u}\rangle - |d\bar{d}\rangle)}{\sqrt{2}}, \quad \eta = \frac{(|u\bar{u}\rangle + |d\bar{d}\rangle - 2\,|s\bar{s}\rangle)}{\sqrt{6}}, \quad \eta' = \frac{(|u\bar{u}\rangle + |d\bar{d}\rangle + |s\bar{s}\rangle)}{\sqrt{3}}. \tag{5.3}$$

It follows that, ignoring factors such as mass difference which break SU(3) isospin symmetry, the $|d\bar{d}\rangle$ state will be found in the ratio (π^0:η:η')=(3:1:2) and thus the neutral kaon states can be substituted by these mesons in this ratio.

For each signal mode in $B \to X_d^H\gamma$ MC data a weight can now be calculated from the known $B \to X_s^H\gamma$ weights. In generated $B \to X_s^H\gamma$ signal MC data after weights are applied, a given mode i will contribute a fraction f_i^s to the total number of events with hadronic mass greater than $1.0\,\mathrm{GeV}/c^2$. Adding the f_i^s of all $B \to X_s^H\gamma$ modes which, on substitution of the kaon, correspond to the signal $B \to X_d^H\gamma$ mode j under

s State	d State	Meson Substitution
$\|s\overline{u}\rangle$	$\|d\overline{u}\rangle$	$K^- \to \pi^-$
$\|s\overline{d}\rangle$	$\|d\overline{d}\rangle$	$\overline{K}^0 \to \pi^0, \eta, \eta'$

TABLE 5.11: Comparison of final state particles when substituting the s quark for a d quark.

consideration gives the expected fraction of mode j events with hadronic mass greater than $1.0\,\mathrm{GeV}/c^2$, f_j^d. The weight, w_j, to apply to $B \to X_d^H\gamma$ events of mode j is then the ratio of this fraction with the original fraction of such events, $f_j^{d'}$, generated in the MC data:

$$w_j = \frac{f_j^d}{f_j^{d'}} = \frac{\sum_i f_i^s}{f_j^{d'}}. \tag{5.4}$$

Table 5.12 lists all of the $B \to X_s^H\gamma$ modes corresponding to a $B \to X_d^H\gamma$ signal mode from kaon substitution as well as showing the calculated weights. The mode $B^+ \to \pi^+\pi^-\pi^+\pi^0\gamma$ can contain contributions from intermediate $\eta \to \pi^+\pi^-\pi^0$ transitions hence under substitution the decay $B^+ \to K^0\pi^+\gamma$ also contributes partially. This is accounted for in the calculation of the weight for this mode using the world average rate of $\eta \to \pi^+\pi^-\pi^0$ decays as a proportion of all potential η meson decays [14].

After weights are applied to the $B \to X_d^H\gamma$ MC data the signal efficiency, ϵ_{MC}, is calculated to be 1.54% which is a variation of 12.0% relative to the default efficiency used to calculate the measured BF. This is therefore assigned as the estimate of systematic uncertainty due to potential differences in MC and experimental data final state distributions for $B \to X_d^H\gamma$ data.

5.1.4.3 Resonant Final States

Uncertainties in the calculated signal efficiency for measurements in the low mass region exist due to the assumed relative contributions of different resonant states. In calculating ϵ_{MC} the relative resonant contributions are assumed to reflect the best measured values taken from [14]. Each measurement has a quoted error which results in a corresponding uncertainty due to that resonance in the efficiency calculation. To numerically evaluate this uncertainty the efficiency is recalculated varying the relative contribution of each resonance within the extremes of its error while leaving the other resonant contributions unchanged. The average relative variation in the efficiency is then interpreted as the uncertainty due to that resonance. Each resonance final state is assumed to have been measured independently thus the individual uncertainties are combined in quadrature to give an overall uncertainty due to all contributing resonances. Tables 5.13 lists the measured rates and associated uncertainties in the calculated signal efficiency of

$B \to X_d^H \gamma$ Mode	Contributing $B \to X_s^H \gamma$ Modes	Calculated Weight
$B^0 \to \pi^+\pi^-\gamma$	$B^0 \to K^+\pi^-\gamma$	0.73
$B^+ \to \pi^+\pi^0\gamma$	$B^0 \to K^+\pi^0\gamma$	
	$B^+ \to K^0\pi^+\gamma$	0.47
$B^+ \to \pi^+\pi^-\pi^+\gamma$	$B^+ \to K^+\pi^-\pi^+\gamma$	1.78
$B^0 \to \pi^+\pi^-\pi^0\gamma$	$B^0 \to K^+\pi^-\pi^0\gamma$	
	$B^0 \to K^0\pi^+\pi^-\gamma$	1.38
$B^0 \to \pi^+\pi^-\pi^+\pi^-\gamma$	$B^0 \to K^+\pi^-\pi^+\pi^-\gamma$	0.77
$B^+ \to \pi^+\pi^-\pi^+\pi^0\gamma$	$B^0 \to K^+\pi^-\pi^+\pi^0\gamma$	
	$B^+ \to K^0\pi^+\pi^-\pi^+\gamma$	
	$B^+ \to K^0\pi^+\gamma$	0.96
$B^+ \to \pi^+\eta\gamma$	$B^+ \to K^+\eta\gamma$	
	$B^+ \to K^0\pi^+\gamma$	1.09

TABLE 5.12: Weights assigned to $B \to X_d^H \gamma$ signal modes to obtain an alternative model of final state contributions. The $B \to X_s^H \gamma$ modes used to calculate each weight, as described in the text, are also shown.

Resonant Decay	BF from [14]	Uncertainty
$B \to K^{*+}\gamma$	$(4.0\pm0.3)\times10^{-5}$	0.4%
$B \to K^{*0}\gamma$	$(4.0\pm0.2)\times10^{-5}$	0.2%
$B \to \rho^+\gamma$	$(8.8^{+2.9}_{-2.5}) \times 10^{-7}$	1.5%
$B \to \rho^0\gamma$	$(9.3\pm2.1)\times10^{-7}$	3.1%
$B \to \omega\gamma$	$(4.6^{+2.0}_{-1.7}) \times 10^{-7}$	7.4%

TABLE 5.13: Uncertainty in signal efficiency due to assumed relative contribution of individual resonant final states.

resonances contributing to $B \to X_s^L \gamma$ and $B \to X_d^L \gamma$ data. The relative contribution of resonant decays to non-reconstructed final states and the proportion of events with hadronic mass less than $1.0\,\mathrm{GeV}/c^2$ are assumed to be accurately modelled by MC data. The combined uncertainty in the signal efficiency is 0.4% for $B \to X_s^L \gamma$ data and 8.2% for the $B \to X_d^L \gamma$ data.

5.1.5 Summary of Experimental Systematic Uncertainties

Table 5.14 summarises all systematic uncertainty estimates detailed above. These are combined in quadrature to give an overall systematic uncertainty for each of the five sources of systematic error outlined and hence a total systematic uncertainty for each measured BF. Table 5.15 summarises the four measured BFs showing the statistical uncertainty and applying the corresponding systematic uncertainties given in table 5.14.

The $B \to X_s^L\gamma$ systematic uncertainty is dominated by the event reductions cuts, in particular the photon selection, PID requirements and NN cut. There is no dominant systematic uncertainty for $B \to X_s^H\gamma$ decays; the event selection, fit and signal model uncertainties all have similar contributions. For both high and low mass $B \to X_d\gamma$ the fit uncertainties, and signal model uncertainties, become more dominant. However, both of these measurements are still statistically limited.

5.2 Measured BF Extrapolation and Extraction of $|V_{td}/V_{ts}|$

Extraction of $|V_{td}/V_{ts}|$ requires extrapolation of the measured BFs to the full $B \to X_{(s,d)}\gamma$ branching ratios required to calculate R in (1.67). This extrapolation is based on the MC data and hence further systematic uncertainties due to potential differences compared to the experimental data must be evaluated. This section describes each BF extrapolation, estimates the associated systematic uncertainties and calculates R to extract $|V_{td}/V_{ts}|$.

5.2.1 Low Mass Region BF Extrapolation

Measured BFs with hadronic mass less than $1.0\,\mathrm{GeV}/c^2$ are modelled entirely as exclusive resonant transitions. Corrections for unreconstructed final states of the subsequent vector meson decays are therefore required.

For $B \to X_s^L\gamma$ data the contribution in MC data from unreconstructed final states are assumed to be accurately represented by the MC data which considers strong decay contributions of K^* mesons as well as $K^* \to K\gamma$ final states. Tables B.2 and B.1 of appendix B show the relative contribution of unreconstructed final states to these transitions with hadronic mass less than $1.0\,\mathrm{GeV}/c^2$ in MC data. In total 50.1% of the final states are not reconstructed. Correcting for this gives an extrapolated BF of $[3.68 \pm 0.14(stat.) \pm 0.20(sys.)] \times 10^{-5}$. As the decays of K^* mesons are well known [14] no systematic uncertainty is applied due to this extrapolation.

The MC model of $B \to X_d^L\gamma$ data only considers strong decays of ρ mesons which correspond to reconstructed modes. The decay model of ω mesons does contain contributions from unreconstructed final states, a summary of which is given table B.3 of appendix B for MC data with hadronic mass less than $1.0\,\mathrm{GeV}/c^2$. In total, after weighting the relative contributions of $(\rho^+{:}\rho^0{:}\omega)$ to world average measurements, 2.0% of final states are not reconstructed. Correcting for these gives an extrapolated BF of $[0.99 \pm 0.27(stat.) \pm 0.14(sys.)] \times 10^{-6}$. The dominant $\omega \to \pi^+\pi^-\pi^0$ transition has an associated uncertainty of 0.8% [14]; however, propagating this as a systematic error in the

Uncertainty Class	Uncertainty Type	$B \to X_s^L \gamma$	$B \to X_s^H \gamma$	$B \to X_d^L \gamma$	$B \to X_d^H \gamma$
Reconstruction Efficiency	Charged Tracks	0.3%	0.5%	0.3%	0.4%
	Neutral Particles	1.6%	1.7%	2.0%	2.2%
	combined	1.6%	1.8%	2.0%	2.2%
Event Reduction	π^0/η Veto	2.0%	2.0%	2.0%	2.0%
	Photon Quality	2.9%	2.9%	2.9%	2.9%
	PID	2.4%	3.5%	2.8%	3.5%
	K_S Veto	-	-	0.2%	0.4%
	Vertex χ^2	0.5%	1.0%	0.5%	1.0%
	Daughter $\lvert p_3 \rvert$	1.0%	1.7%	1.0%	1.7%
	Neural Net	2.5%	3.1%	2.5%	3.1%
	combined	5.1%	6.2%	5.3%	6.2%
EMLF	PDF Parameters	0.3%	2.9%	5.9%	3.6%
	Fixed Yields	2.3%	4.5%	8.8%	9.1%
	Fit Bias	0.6%	0.9%	2.9%	1.3%
	combined	2.4%	5.4%	11.0%	9.9%
Signal Model	JETSET phase space	-	3.5%	-	12.0%
	KN Spectrum	-	5.8%	-	6.0%
	PDG Uncertainty	0.4%	-	8.2%	-
	combined	0.4%	6.8%	8.2%	13.4%
B Counting		0.6%	0.6%	0.6%	0.6%
	Total Systematic Uncertainty	5.9%	10.8%	14.9%	17.9%

TABLE 5.14: Estimates of the total systematic uncertainty for each measured BF.

Fit	Measured BF
$B \to X_s^L \gamma$	$(1.84 \pm 0.07 \pm 0.11) \times 10^{-5}$
$B \to X_s^H \gamma$	$(6.30 \pm 0.24 \pm 0.68) \times 10^{-5}$
$B \to X_d^L \gamma$	$(0.97 \pm 0.27 \pm 0.14) \times 10^{-6}$
$B \to X_d^H \gamma$	$(3.60 \pm 0.92 \pm 0.64) \times 10^{-6}$

TABLE 5.15: Measured BF estimation showing associated statistical error for each fit to experimental data and the corresponding estimate of systematic uncertainty before BF extrapolation.

extrapolation shows it has a negligible contribution to the overall systematic uncertainty and so it is not considered further.

5.2.2 High Mass Region BF Extrapolation

Measured BFs with hadronic mass in the range $1.0 \leq m_X \leq 2.0\,\mathrm{GeV}/c^2$ are used to estimate the total BF of $B \to X_{(s,d)}\gamma$ transitions with mass greater than $1.0\,\mathrm{GeV}/c^2$. This extrapolation is made in two steps, firstly a total BF in the measured mass range is estimated and then the KN model is used to infer the correction required for unreconstructed states with mass greater than $2.0\,\mathrm{GeV}/c^2$.

5.2.2.1 Correction for Final States with $1.0 \leq m_X \leq 2.0\,\mathrm{GeV}/c^2$

The total number of events in the mass range $1.0 \leq m_X \leq 2.0\,\mathrm{GeV}/c^2$ is determined by the KN model parameterisation. With the default MC models (weighted for $B \to X_s^H \gamma$ data and as generated for $B \to X_d^H \gamma$ data) the correction is simply determined by the ratio of the total number of signal events to all events in the mass region. For $B \to X_s^H \gamma$ MC data 35.4% of events in this mass range are signal modes giving an extrapolated BF of $[17.8 \pm 0.7(stat.) \pm 1.9(sys.)] \times 10^{-5}$. The corresponding signal proportion in $B \to X_d^H \gamma$ MC data is 42.3% of events in the mass range which gives extrapolated BF $[8.5 \pm 2.2(stat.) \pm 1.5(sys.)] \times 10^{-6}$. However, the proportion of signal events in the mass range is not necessarily well modelled due to uncertainties in the JETSET phase space decay model so further systematic uncertainties must be applied to the extrapolated BFs to account for this.

In section 5.1.4.2 varying the relative contributions of signal modes was used to determine a systematic uncertainty in the signal efficiency. In these alternative models the total number of signal events were not constant. It follows that, assuming the model changes do not significantly alter the results of fits to data, these variations result in a different extrapolated BF as well as a different measured BF. The errors must therefore

Signal Mode	Uncertainty
$B^0 \to K^+\pi^-\gamma$	1.8%
$B^0 \to K^+\pi^0\gamma$	1.2%
$B^0 \to K^+\pi^-\pi^+\gamma$	2.9%
$B^0 \to K^+\pi^-\pi^0\gamma$	1.5%
$B^0 \to K^+\pi^-\pi^+\pi^-\gamma$	0.6%
$B^0 \to K^+\pi^-\pi^+\pi^0\gamma$	0.9%
$B^0 \to K^+\eta\gamma$	1.4%
All Signal	4.3%

TABLE 5.16: Systematic uncertainties due to uncertainties in the weights applied to $B \to X_s^H\gamma$ signal MC data modes after BF extrapolation in the signal mass range. These values modify those given in table 5.10 as potentially correlated variations in the overall proportion of signal events within the mass range are accounted for.

be propagated for the extrapolated BF accounting for this effect. Table 5.16 lists the average variation in the extrapolated BFs obtained from the existing calculated uncertainty for each $B \to X_s^H\gamma$ signal mode. These are then combined in quadrature to give an overall corrected systematic uncertainty of 4.3% for the extrapolated BF. When propagating the alternative $B \to X_d^H\gamma$ signal model systematic uncertainty, the relative proportion of signal events in the new model is 45.0% and the measured BF due to the efficiency change is 4.09×10^{-6}. This therefore results in an extrapolated BF which varies by 6.9% from the default value. This is assigned as the corrected systematic for the extrapolated BF of these data.

Further uncertainties in the proportion of signal events within the signal mass region arise from the modelled distribution of unreconstructed modes the relative contributions of which are not well known. These are classified into two categories, those with 2-4 bodies in the final state and those with 5 or more bodies in the final state. The previous version of this analysis investigated the contribution from such states by engineering alternative fragmentation models of the hadronic system and seeing how the proportion of those unreconstructed classes change relative to the default model [1]. Models considered are applied to both X_s and X_d MC data and include the default JETSET fragmentation, the application of weights measured in previous $B \to X_s\gamma$ experimental datasets (see section 4.2) and a hybrid model of exclusive resonant MC data with the JETSET events. The hybrid model uses the resonant decays listed in table 5.17; the resonant decays are assigned half of the the total inclusive width in this region with the other half remain JETSET decays. Each resonance is assumed to contribute equally. The variations in unknown modes under these alternative models corresponds to no more than 50% for either category, hence this can be considered a conservative limit. The previous analysis study motivates assigning a systematic uncertainty on the extrapolated

$B \to X_s^H \gamma$	$B \to X_d^H \gamma$
$B \to K_1(1270)\gamma$	$B \to h_1(1170)\gamma$
$B \to K_1(1400)\gamma$	$B \to b_1(1235)\gamma$
$B \to K^*(1410)\gamma$	$B \to a_1(1260)\gamma$
$B \to K_2^*(1430)\gamma$	$B \to f_2^0(1270)\gamma$
$B \to K^*(1680)\gamma$	$B \to f_1^0(1285)\gamma$
	$B \to a_2^0(1320)\gamma$

TABLE 5.17: Resonant decays used to construct the hybrid fragmentation model investigated in the previous version of this analysis [1].

BF whereby unreconstructed contributions in the default models are varied individually by ±50%. The uncertainty for a given unknown category is calculated by varying one unknown mode while keeping the relative contribution of signal and the remaining unknown mode constant, but renormalising them relative to the varied mode to maintain overall unity in the mass range. The resulting variation in the signal contribution then forms the systematic uncertainty; if the variation is asymmetric the larger value is taken. By keeping the other unknown mode constant relative to the signal in a given variation the uncertainties from each unknown class are to a good approximation independent and hence their sum in quadrature gives the overall uncertainty due to unreconstructed final states from JETSET decays.

The distribution of unreconstructed states in $B \to X_s^H \gamma$ MC data do not include decays corresponding to a signal mode where the K^+ final state meson is substituted for a K^0 meson. This is because such decays exhibit an isospin symmetry with the signal modes and are therefore constrained by that symmetry. In the signal mass range the proportion of signal modes is 35.4%, corresponding K^0 modes 36.4%, other 2-4 body modes 16.0% and 5 or more body modes 12.2%. Varying the 2-4 body mode proportion by ±50% gives an average relative change in the extrapolated BF of 9.6%. The corresponding variation of 5 or more body modes gives a change of 7.0%. An overall systematic uncertainty due to unreconstructed modes is therefore taken to be 11.9%. Accounting for this and the above corrected signal JETSET uncertainty the extrapolated BF in the signal mass range for $B \to X_s^H \gamma$ data is therefore $[17.8 \pm 0.7(stat.) \pm 2.9(sys.)] \times 10^{-5}$.

Unreconstructed states in $B \to X_d^H \gamma$ MC data are all varied as none are expected to be constrained by the signal modes. In the signal mass range for these data the proportion of signal modes is 42.3%, other 2-4 body modes 27.0% and 5 or more body modes 30.7%. The uncertainty obtained from varying the 2-4 body mode proportion by ±50% is 19.2% and the corresponding uncertainty due to 5 or more body modes is 23.3%. This gives an overall systematic uncertainty due to unreconstructed modes of 30.2%. Correcting the systematic error to account for the above extrapolation uncertainties gives

the extrapolated BF in the signal mass range for $B \to X_d^H \gamma$ data as $[8.5 \pm 2.2(stat.) \pm 2.9(sys.)] \times 10^{-6}$.

5.2.2.2 Correction for Final States with m_X >2.0 GeV/c^2

Given the extrapolated BF values calculated in the hadronic mass range $1.0 \leq m_X \leq 2.0\,\mathrm{GeV}/c^2$, estimates of the total BF for all masses greater than $1.0\,\mathrm{GeV}/c^2$ are made for both $B \to X_s^H \gamma$ and $B \to X_d^H \gamma$ data. The generated signal MC data using the default KN model parameterisation shows that for all masses greater than $1.0\,\mathrm{GeV}/c^2$ in $B \to X_s^H \gamma$ data, 60.2% of events are generated with $1.0 \leq m_X \leq 2.0\,\mathrm{GeV}/c^2$. Given this proportion the total BF with mass greater than $1.0\,\mathrm{GeV}/c^2$ is calculated to be $[29.5 \pm 1.1(stat.) \pm 4.8(sys.)] \times 10^{-5}$. For $B \to X_d^H \gamma$ data the corresponding proportion of events with $1.0 \leq m_X \leq 2.0\,\mathrm{GeV}/c^2$ is also 60.2% which gives the total BF with mass greater than $1.0\,\mathrm{GeV}/c^2$ as $[14.2 \pm 3.6(stat.) \pm 4.8(sys.)] \times 10^{-6}$. However, the KN model contribution to the systematic uncertainties in the above BFs must be corrected to account for uncertainties in the extrapolation due to the parameterisation of this model. For each transition flavour the complete high mass extrapolation is repeated assuming alternative KN parameterisations. The efficiencies given in table 5.8 are used to revise the measured BF calculation assuming no change in the result of the fit to experimental data. The BFs for all hadronic masses greater than $1.0\,\mathrm{GeV}/c^2$ are then recalculated using the relevant proportions of signal candidates under the alternative KN scheme. The average variation in the full BF for $B \to X_s^H \gamma$ data is 1.6% and the corresponding variation in $B \to X_d^H \gamma$ data is 1.9%. These are therefore assigned as the corrected systematic uncertainties due to different parameterisations of the KN model. With this correction applied the the total BFs for transitions with hadronic mass greater than $1.0\,\mathrm{GeV}/c^2$ is $[29.5 \pm 1.1(stat.) \pm 4.5(sys.)] \times 10^{-5}$ for $B \to X_s^H \gamma$ data and $[14.2 \pm 3.6(stat.) \pm 4.7(sys.)] \times 10^{-6}$ for $B \to X_d^H \gamma$ data.

5.2.2.3 Summary of Extrapolated BFs

Table 5.18 summarises the extrapolated BFs calculated for each mass region. For the high mass region both the extrapolated BF in the hadronic mass range $1.0 \leq m_X \leq 2.0\,\mathrm{GeV}/c^2$ and the extrapolated BF for all hadronic masses with $1.0\,\mathrm{GeV}/c^2 \leq m_X$ are shown. The full extrapolated BFs in the high and low mass regions are not directly combined into an overall inclusive BF due to correlations in the systematic uncertainties which are discussed below.

Fit	Hadronic Mass Range	Extrapolated BF
$B \to X_s^L \gamma$	$<1.0\,\mathrm{GeV}/c^2$	$(3.68\pm0.14\pm0.20)\times10^{-5}$
$B \to X_s^H \gamma$	1.0-2.0 GeV/c^2	$(17.8\pm0.7\pm2.9)\times10^{-5}$
	$\geq1.0\,\mathrm{GeV}/c^2$	$(29.5\pm1.1\pm4.5)\times10^{-5}$
$B \to X_d^L \gamma$	$<1.0\,\mathrm{GeV}/c^2$	$(0.99\pm0.27\pm0.14)\times10^{-6}$
$B \to X_d^H \gamma$	1.0-2.0 GeV/c^2	$(8.5\pm2.2\pm2.9)\times10^{-6}$
	$\geq1.0\,\mathrm{GeV}/c^2$	$(14.2\pm3.6\pm4.7)\times10^{-6}$

TABLE 5.18: Extrapolated BF estimation showing the associated statistical error for each fit to experimental data and the corresponding estimate of systematic uncertainty which has been recalculated to account for uncertainties in the extrapolation method.

5.2.3 Calculation of $|V_{td}/V_{ts}|$ and Associated Uncertainty

The evaluation of R is given by the relation,

$$R = \frac{\mathcal{B}(B \to X_d\gamma)^{m_x<1.0\,\mathrm{GeV}/c^2} + \mathcal{B}(B \to X_d\gamma)^{m_x>1.0\,\mathrm{GeV}/c^2}}{\mathcal{B}(B \to X_s\gamma)^{m_x<1.0\,\mathrm{GeV}/c^2} + \mathcal{B}(B \to X_s\gamma)^{m_x>1.0\,\mathrm{GeV}/c^2}}. \tag{5.5}$$

Using the extrapolated BFs calculated above therefore gives $R = (15.14\times10^{-6})/(33.22\times10^{-5}) = 0.0456$. The corresponding uncertainties in R must take account any expected correlations in the contributions to the uncertainties on each extrapolated BF. There are four possible correlations for each uncertainty. Firstly no correlation in the uncertainty; in this case the average variation in R must be calculated separately for the variation due to each individual BF and these contributions then added in quadrature. Secondly the BFs corresponding to the same flavour transition have a correlated uncertainty; in this case the the average variation in R due to each flavour transition is calculated and these are added in quadrature. Thirdly the BFs corresponding to the same mass range have a correlated uncertainty; similarly the average variation in R due to each mass range is calculated and these are added in quadrature. Finally the uncertainty is correlated for all BFs; in this case the average variation in R due to all BFs is calculated. The propagation of the overall PID systematic is calculated as the linear combination of the individual charged pion and charged kaon PID contributions to the uncertainty in R. This is due to potential correlations in the input variables for each PID selector. The charged pion PID uncertainty in R is correlated for all BFs and calculated to be 1.2%. The charged kaon PID contribution is also correlated for all BFs (although $B \to X_d\gamma$ BFs have no contribution from this uncertainty) and calculated to give a 1.1% uncertainty in R. Table 5.19 list all sources of uncertainty calculated for the individual BFs, indicates their expected correlation and evaluates the corresponding uncertainty on R. Combining

Uncertainty Type	Correlation Type	R Uncertainty
Fit; Statistical	None	24.2%
Track Reconstruction	All	0.1%
Neutral Reconstruction	All	0.5%
Photon Cuts	All	0.0%
Combined PID	All	2.3%
K_S Veto	Flavour	0.4%
Vertex χ^2	Mass Range	0.1%
Daughter $\|p_3\|$	Mass Range	0.1%
Neural Net	Mass Range	0.1%
Fit; PDF Fixed Parameters	None	4.3%
Fit; Fixed $B \to X_s^L\gamma$ XF Yield	All	0.5%
Fit; Fixed $B \to X_s^H\gamma$ XF Yield	All	0.0%
Fit; Fixed $B \to X_d^L\gamma$ XF Yield	Flavour	0.6%
Fit; Fixed $B \to X_d^H\gamma$ XF Yield	Flavour	1.9%
Fit; Fixed PID XF Yield	Flavour	1.2%
Fit; Fixed Generic B Yields	All	5.6%
Fit; Floated Parameter Bias	None	1.5%
JETSET Model Signal Modes	None	7.5%
JETSET Model Unreconstructed 2-4 Body Modes	Mass Range	9.5%
JETSET Model Unreconstructed 5+ Body Modes	Mass Range	15.6%
KN Spectrum	Mass Range	0.4%
PDG Uncertainty	None	0.5%
B Counting	All	0.0%
Combined Systematic	-	21.3%

TABLE 5.19: Estimated uncertainties in the calculation of R. There are four correlation types. 'All' represents the case where all four extrapolated BFs are varied simultaneously. 'Flavour' represents the case where BFs of the same flavour are varied simultaneously and the resulting uncertainties in R combined in quadrature. 'Mass Range' represents the case where BFs in the same mass region are varied simultaneously and the resulting uncertainties in R combined in quadrature. 'None' represents the case where each BF is varied independently and the resulting uncertainties in R combined in quadrature.

these individual uncertainty contributions in quadrature gives,

$$R = 0.0456 \pm 0.0110(stat.) \pm 0.0097(sys.). \tag{5.6}$$

Solving (1.67) for the calculated value of R gives,

$$\left|\frac{V_{td}}{V_{ts}}\right| = 0.211 \pm 0.023(stat.) \pm 0.022(sys.) \pm 0.001(th.), \tag{5.7}$$

where the final uncertainty is the associated theoretical error estimated numerically from the variations shown in figure 1.9. This value is in good agreement with the world average value from B mixing experiments which is 0.209 ± 0.006 [14]. The systematic uncertainty of this measurement is dominated by the uncertainties due to the JETSET phase space model, particularly in extrapolating the measured BFs to inclusive values. The systematic and statistical uncertainties are of a similar order.

Chapter 6

Summary and Future Prospects

The analysis presented uses experimental data recorded with the *BABAR* detector to measure the sum of seven exclusive $B \to X_s\gamma$ and $B \to X_d\gamma$ decay BFs in the hadronic mass ranges $0.5 \leq m_X < 1.0\,\mathrm{GeV}/c^2$ and $1.0 \leq m_X \leq 2.0\,\mathrm{GeV}/c^2$. For each mass range the event selection has been optimised with simulated data to maximise the signal yield of CKM-suppressed $B \to X_d\gamma$ transitions over their associated combinatoric backgrounds. Identical event selection has been applied to $B \to X_s\gamma$ transitions as to their $B \to X_d\gamma$ counterparts. Signal yields in experimental data have been extracted with unbinned extended maximum likelihood fits and interpreted as BF measurements. Comprehensive evaluation of the systematic uncertainties relating to these measurements have been made. Table 6.1 lists the measured BF for each transition flavour in each mass range and their associated uncertainties. The measured $B \to X_d\gamma$ BF in the mass range $1.0 \leq m_X \leq 2.0\,\mathrm{GeV}/c^2$ represents the first statistically significant signal yield of such transitions with hadronic mass above the ρ/ω resonant peaks.

The measured BFs have been extrapolated in a model dependent way to estimate BFs for all $B \to X_s\gamma$ and $B \to X_d\gamma$ transitions in the hadronic mass ranges $m_X < 1.0\,\mathrm{GeV}/c^2$ and $m_X \geq 1.0\,\mathrm{GeV}/c^2$. Evaluations of the associated systematic uncertainties due to the extrapolation technique have been made. Table 6.2 lists the extrapolated BF for each transition flavour in each mass range and their associated uncertainties. The existing world average values of these measurements from [14] are shown in table 6.3. Allowing

	$0.5 \leq m_X < 1.0\,\mathrm{GeV}/c^2$	$1.0 \leq m_X \leq 2.0\,\mathrm{GeV}/c^2$
$B \to X_s\gamma$	$(1.84\pm0.07\pm0.11)\times10^{-5}$	$(6.30\pm0.24\pm0.68)\times10^{-5}$
$B \to X_d\gamma$	$(0.97\pm0.27\pm0.14)\times10^{-6}$	$(3.60\pm0.92\pm0.64)\times10^{-6}$

TABLE 6.1: Summary of measured BFs. In each case the first error is statistical and the second systematic.

	$m_X < 1.0\,\mathrm{GeV}/c^2$	$1.0 \leq m_X\,\mathrm{GeV}/c^2$
$B \to X_s\gamma$	$(3.68\pm0.14\pm0.20)\times10^{-5}$	$(29.5\pm1.1\pm4.5)\times10^{-5}$
$B \to X_d\gamma$	$(0.99\pm0.27\pm0.14)\times10^{-6}$	$(14.2\pm3.6\pm4.7)\times10^{-6}$

TABLE 6.2: Summary of extrapolated BFs. In each case the first error is statistical and the second systematic.

Mode	BF from [14]
$B \to K^*\gamma$	$(4.2 \pm 0.6) \times 10^{-5}$
$\overline{b} \to \overline{s}\gamma$	$(35.6 \pm 2.5) \times 10^{-5}$
$B \to \rho/\omega\gamma$	$(1.3 \pm 0.2) \times 10^{-6}$

TABLE 6.3: Existing world average BF measurements.

Measurement	$\lvert V_{td}/V_{ts}\rvert$
Inclusive Penguin	0.211±0.032
Exclusive Penguin [14]	0.210±0.040
B Mixing [14]	0.209±0.006

TABLE 6.4: $|V_{td}/V_{ts}|$ measurement comparisons.

for the fact that 10% of K^* and ρ resonance peaks have hadronic mass greater than $1.0\,\mathrm{GeV}/c^2$ is can be seen that the extrapolated BFs are in good agreement with the world average values.

The extrapolated BFs are combined to give a ratio of the $B \to X_d\gamma$ transition to the $B \to X_s\gamma$ transition of $R(d\gamma/s\gamma) = 0.0456 \pm 0.0110(stat.) \pm 0.0097(sys.)$. This parameter is interpreted using its theoretical relation to the CKM parameters $(\overline{\rho}, \overline{\eta})$ presented in [22]. After transforming the basis of this relation to the orthogonal coordinates $(|V_{td}/V_{ts}|, \beta)$, discussed in section 1.2.2.2, the calculated value of $R(d\gamma/s\gamma)$ is found to correspond to $|V_{td}/V_{ts}| = 0.211 \pm 0.023(stat.) \pm 0.022(sys.) \pm 0.001(th.)$ where the final error corresponds to the uncertainty in the theoretical relation. The extracted value of $|V_{td}/V_{ts}|$ is compared to previous measurements in table 6.4; this table essentially summarises the measurements discussed in section 1.3. It is found to be entirely consistent and has uncertainties comparable to those from exclusive radiative penguin measurements; furthermore it makes a significant improvement compared the previous version of this analysis [1]. The agreement of both this and the exclusive radiative penguin measurement with the value extracted from B mixing suggests no significant new physics amplitude due to the operator $C_7(\mu)Q_7$ which exclusively contributes to calculations of the radiative penguin processes.

6.1 Future Prospects

This measurement is performed on the complete *BABAR* experimental dataset collected at the $\Upsilon(4S)$ resonance and so is unlikely to be improved by the *BABAR* collaboration. However, there is an ongoing *BABAR* analysis to update the $B \to X_s\gamma$ measurement from [30] using the full experimental dataset. The systematic uncertainty due to model dependancies in $B \to X_s^H\gamma$ MC data could be significantly reduced using the results of such an analysis, although these do not dominate the total systematic uncertainty calculated here.

Other future prospects for improving the measurement of $|V_{td}/V_{ts}|$ from radiative penguin decays are threefold. Firstly the BELLE experiment at KEK in Japan [72] also has a significant volume of e^+e^- collision data recorded at the $\Upsilon(4S)$ resonance, the collection of which is ongoing at the time of writing. These data could well be used to perform an analysis similar to the one presented here, reducing the statistical uncertainties due to $B \to X_d\gamma$ transitions. However, the measurement is unlikely to improve the modelling uncertainties due to JETSET. Secondly the LHCb experiment at CERN [73] has started to record proton-proton collision data with the aim of measuring processes from both B and B_s meson decays. The noisier environment of hadronic collisions makes the prospect of measuring $|V_{td}/V_{ts}|$ from inclusive decays unlikely, although studies suggest a competitive measurement from exclusive decays is possible and this forms an important part of their physics program [74]. However, such a measurement from exclusive decays will suffer from the larger theoretical uncertainties discussed in section 1.2.2.3. Finally proposals exist for a next generation of high luminosity e^+e^- collision experiments dedicated to data collection at the $\Upsilon(4S)$ resonance [75, 76]. The proposal in [75] uses a detector design heavily based on the *BABAR* experiment and in fact proposes to reuse components of the now dismantled *BABAR* detector. These designs project potential datasets in excess of $30\,\mathrm{ab}^{-1}$ which could potentially make the $B \to X_d\gamma$ transition measurements as competitive as existing $B \to X_s\gamma$ measurements. It follows that a measurement of $|V_{td}/V_{ts}|$ similar to the measurement presented here could be made with much reduced statistical error. If this were to be the case such a measurement would also have to undertake significant studies to reduce the dominant systematic uncertainties. Datasets of this size could well be used to measure a significantly larger proportion of $B \to X_d\gamma$ decay modes at a relatively small cost in statistical power (as demonstrated when modes eight and nine of table 3.1 were investigated here). This would reduce the modelling uncertainty; however the increased contribution of combinatoric backgrounds from generic B decays would also need extensive study as that was shown to be a significant contribution to the systematic uncertainty here. In any case these high luminosity experiments, if approved, are unlikely to begin data collection

before 2015. Significant improvement of the measurement presented here is therefore unlikely in the near future.

6.2 Addendum

The analysis presented in this thesis has been modified by the *BABAR* collaboration, with assistance from the author, and was submitted to Physical Review Letters on 21st May 2010 [77]. The modifications to the methods presented in this thesis are briefly summarised.

Event selection is unmodified and the MC and experimental datasets used in MLFs are identical. The fits to data are significantly modified. The generic B normalisation is allowed to float at a cost in fit bias on the signal yield. However, the additional bias is significantly less than the systematic uncertainties associated with fixed normalisations for this background class which were calculated here. In addition the weights for measured $B \to X_s\gamma$ modes are evaluated from fits to the high mass $B \to X_s\gamma$ experimental data. They are generally found to be consistent with weights from the previous version of this analysis, but do modify the the signal efficiency of the $B \to X_s^H\gamma$ measurement and thus give a different BF; this also modifies the signal model systematic uncertainty calculated when the weights are applied to $B \to X_d^H\gamma$ MC data. Furthermore the submitted analysis assumes isospin holds for ρ and ω resonant decays, whereas the analysis here uses the central value of previous exclusive radiative penguin measurements. This sightly modifies the measured and extrapolated $B \to X_d^L\gamma$ BF. The evaluation of the NN systematic uncertainty is modified but still evaluated with the high statistics $B \to X_s\gamma$ experimental data. The NN systematic is used to correct the efficiency in MC data which is not done here. This results in different signal efficiencies and hence measured BFs for all experimental data. The BF extrapolation is not significantly modified; however, an additional alternative JETSET fragmentation model is engineered whereby only low mass scalar mesons are allowed. This does not increase the $\pm 50\%$ variation observed previously in unreconstructed modes. The published analysis finds $|V_{td}/V_{ts}|$=0.199$\pm$0.022$\pm$0.024$\pm$0.002 [77] where the first error is statistical, the second error systematic and third error theoretical. The theoretical error is slightly larger as it corresponds to a different value of the ratio R. This result varies from the measurement in this thesis by $\sim$5%; the variation is entirely consistent with the systematic difference in the fit method and the modified NN systematic in the published analysis.

Appendix A

Monomial Functions and Sphericity Tensor Variable Distributions

This appendix contains plots of the monomial functions and sphericity tensor variables discussed in sections 3.3.2.2 and 3.3.2.3 respectively. In all cases the solid line shows the distribution for correctly reconstructed signal candidates before event selection cuts are applied and the dashed line shows the corresponding distribution for continuum backgrounds.

A.1 Monomial Function Distributions

Figures A.1-A.3 show the longitudinal monomials for the ROE with respect to the thrust axis of the ROE calculated in the CM frame. Note that L_1^T is identical to the thrust of the ROE. Figures A.4-A.6 show the perpendicular monomials for the ROE with respect to the thrust axis of the ROE calculated in the CM frame. Figures A.7-A.9 show the longitudinal monomials for the ROE with the momentum axis of the high energy photon used in candidate reconstruction calculated in the CM frame. Figures A.10-A.12 show the perpendicular monomials for the ROE with the momentum axis of the high energy photon used in candidate reconstruction calculated in the CM frame.

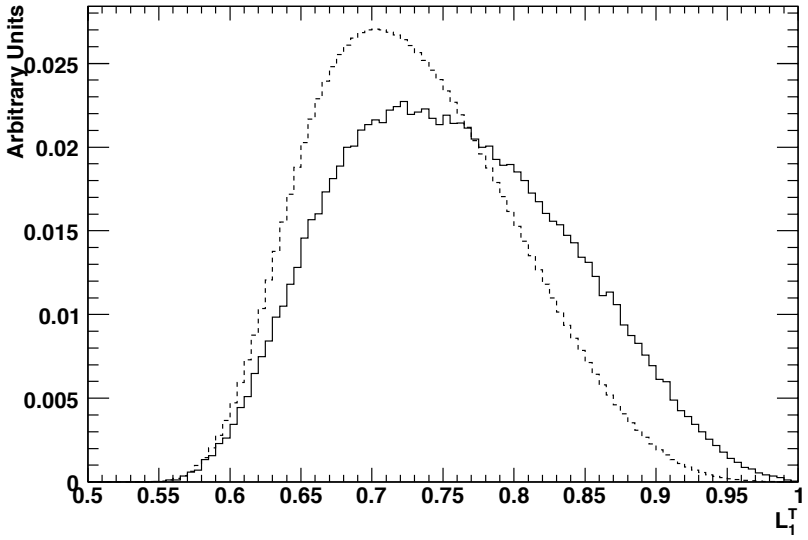


FIGURE A.1: Distributions of the monomial L_1^T

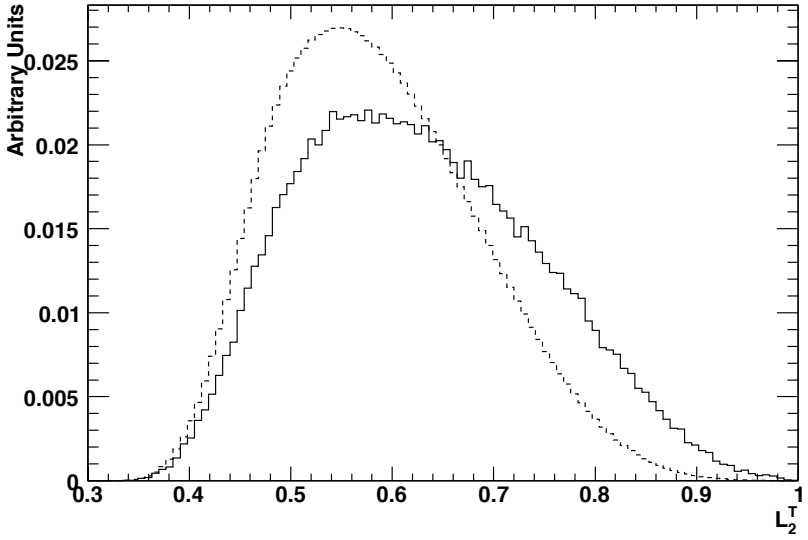


FIGURE A.2: Distributions of the monomial L_2^T

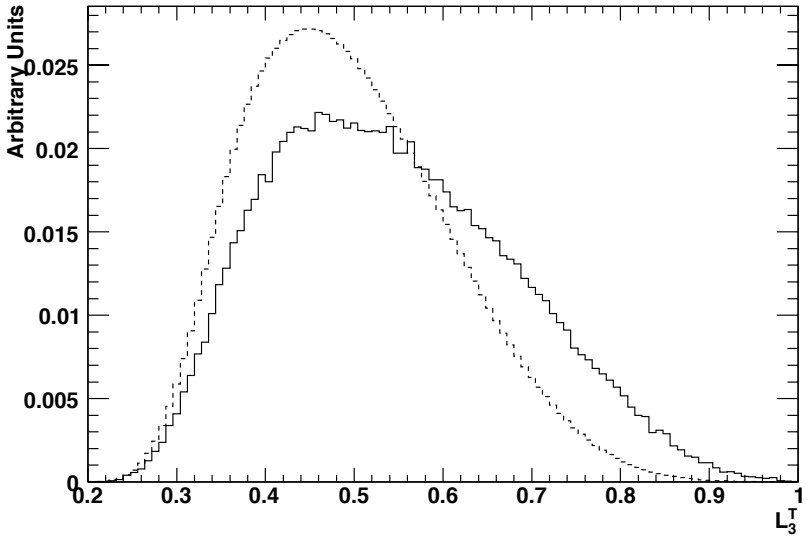


FIGURE A.3: Distributions of the monomial L_3^T

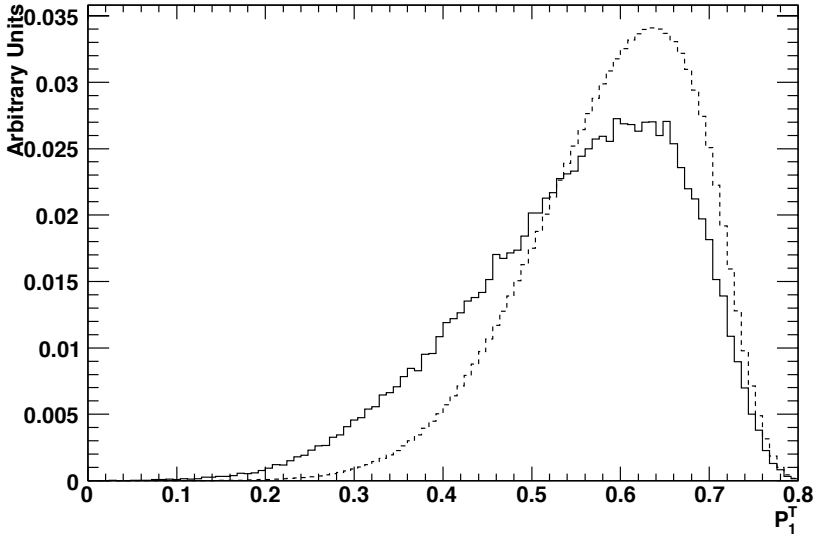


FIGURE A.4: Distributions of the monomial P_1^T

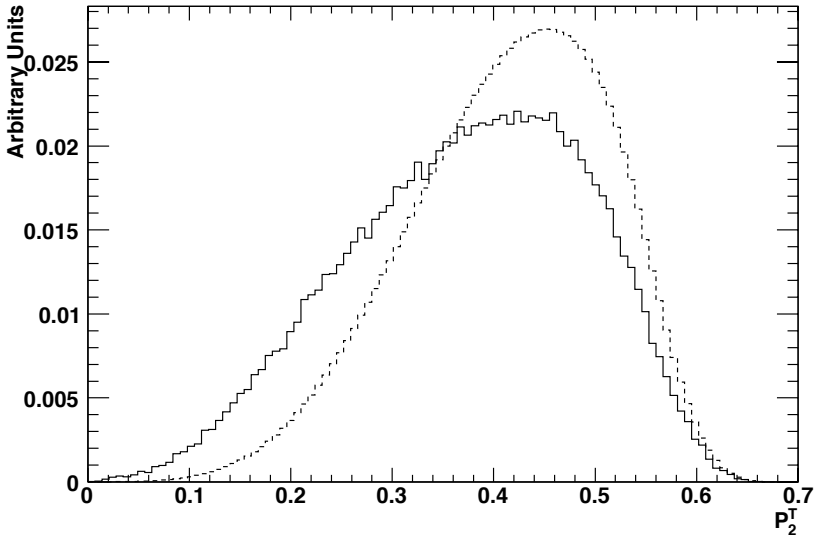


FIGURE A.5: Distributions of the monomial P_2^T

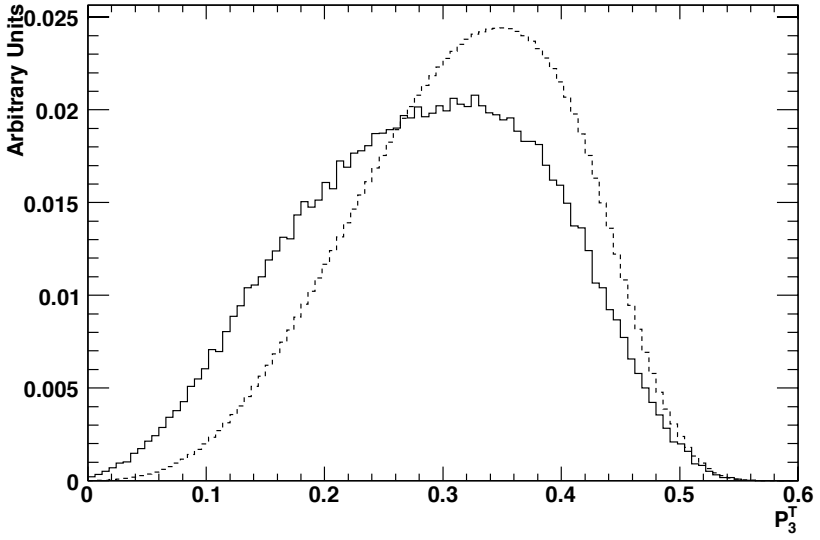


FIGURE A.6: Distributions of the monomial P_3^T

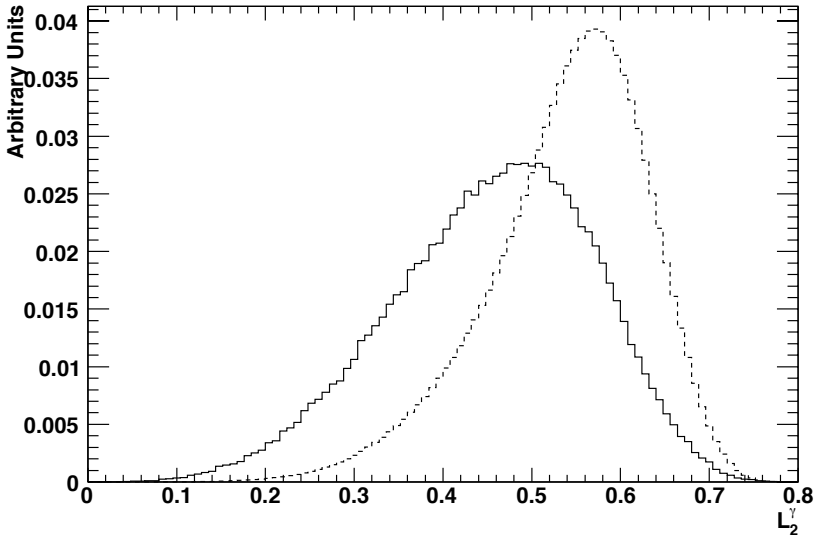


FIGURE A.7: Distributions of the monomial L_1^γ

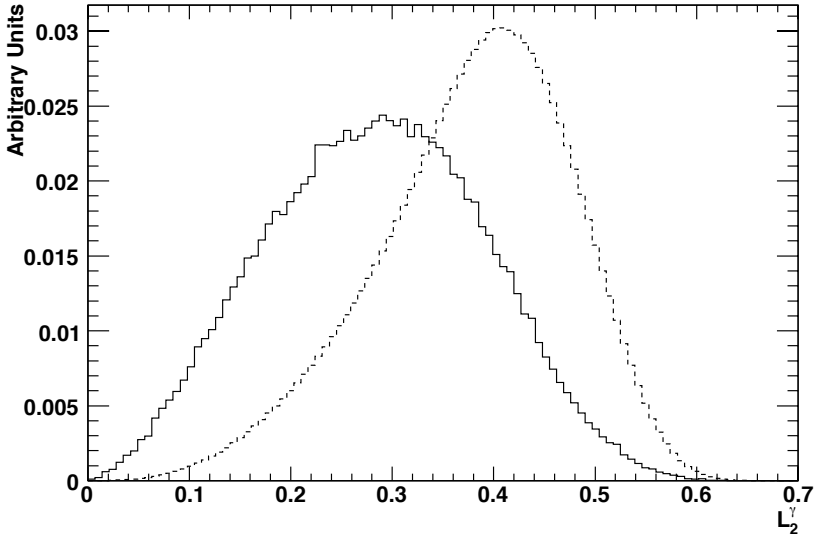


FIGURE A.8: Distributions of the monomial L_2^γ

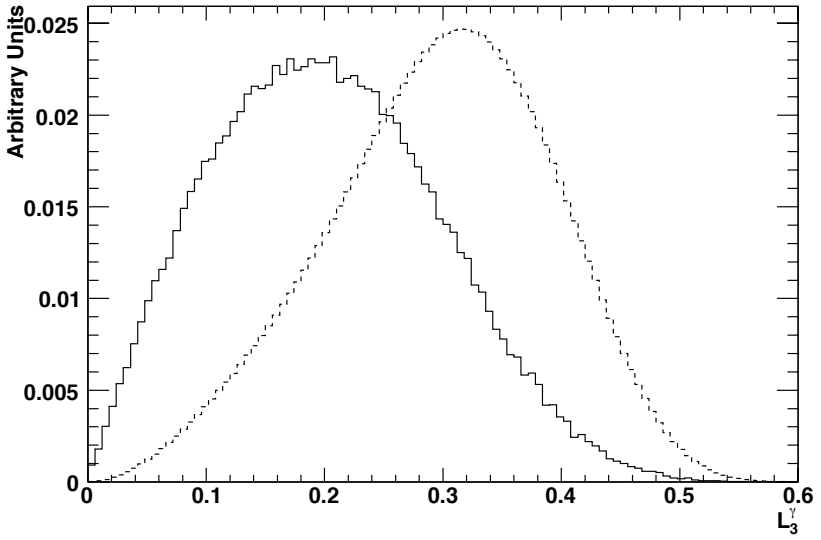


FIGURE A.9: Distributions of the monomial L_3^γ

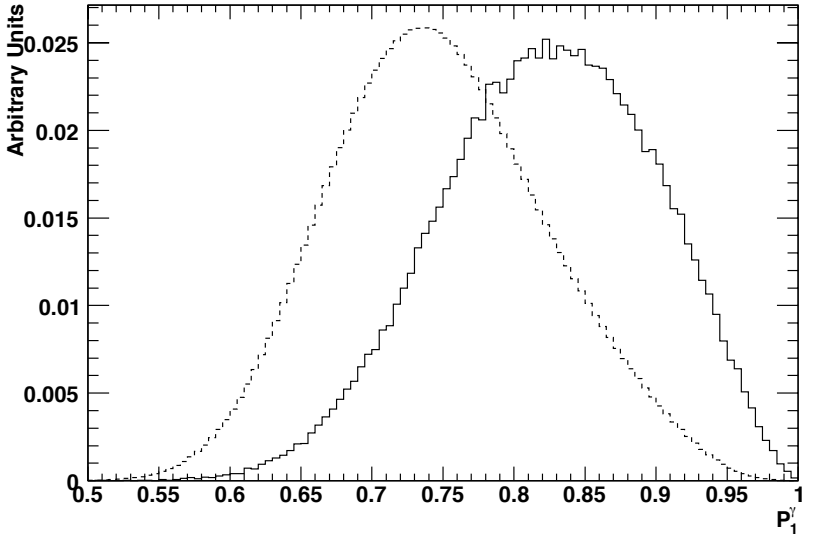


FIGURE A.10: Distributions of the monomial P_1^γ

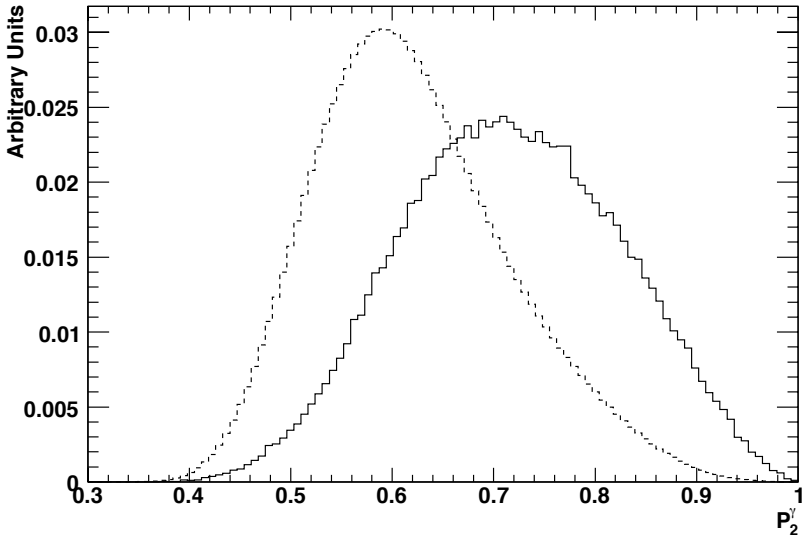


FIGURE A.11: Distributions of the monomial P_2^γ

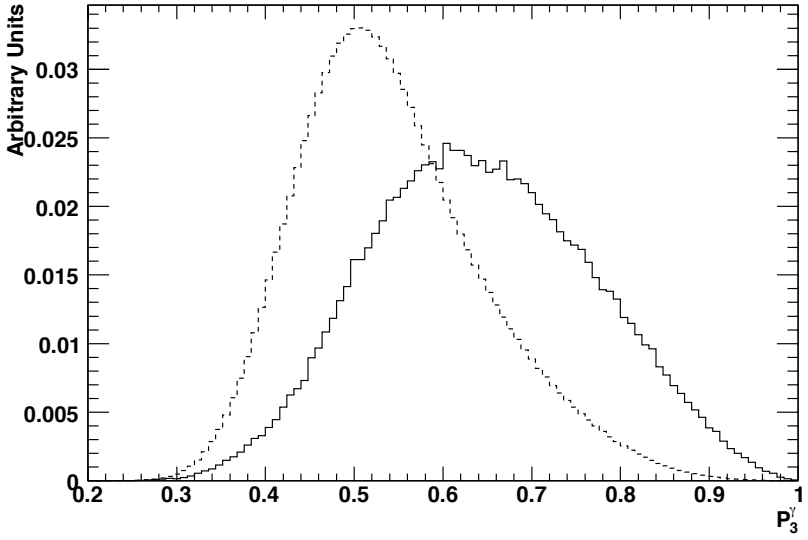


FIGURE A.12: Distributions of the monomial P_2^γ

A.2 Sphericity Tensor Variable Distributions

Figure A.13 shows the sphericity of the ROE, figure A.14 shows the planarity of the ROE and A.15 shows the aplanarity of the ROE.

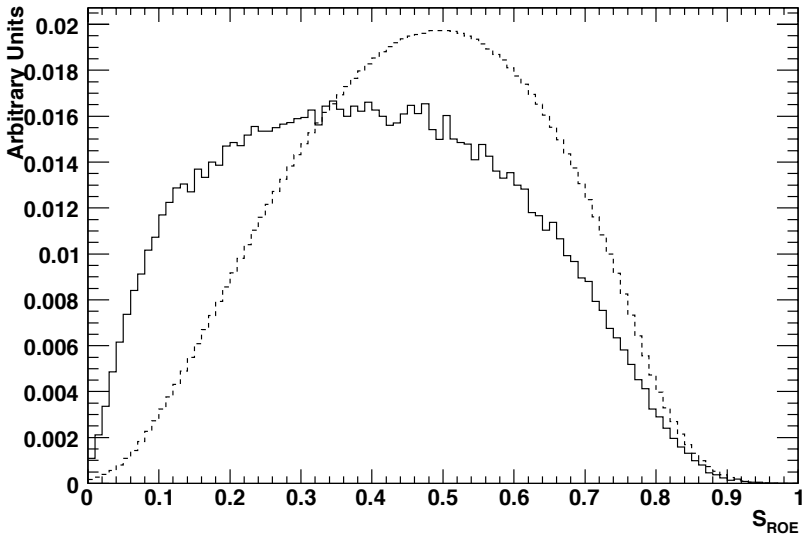


FIGURE A.13: Distributions of S_{ROE}

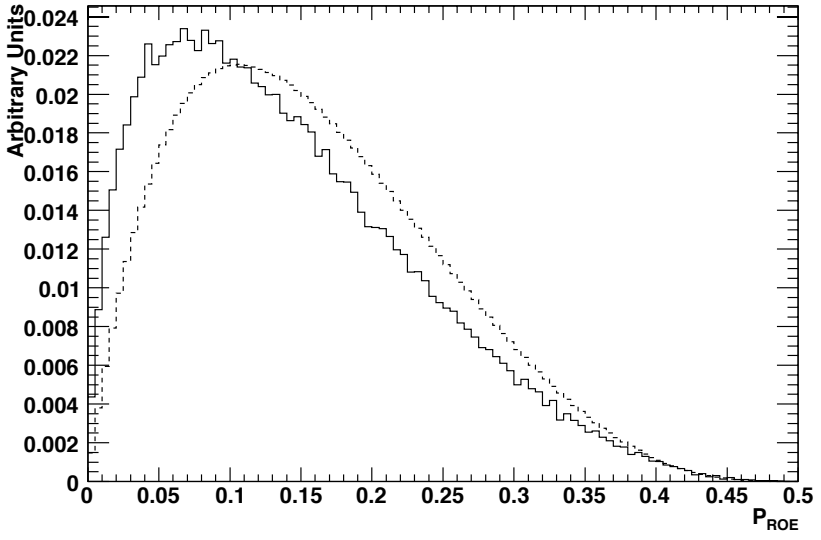


FIGURE A.14: Distributions of P_{ROE}

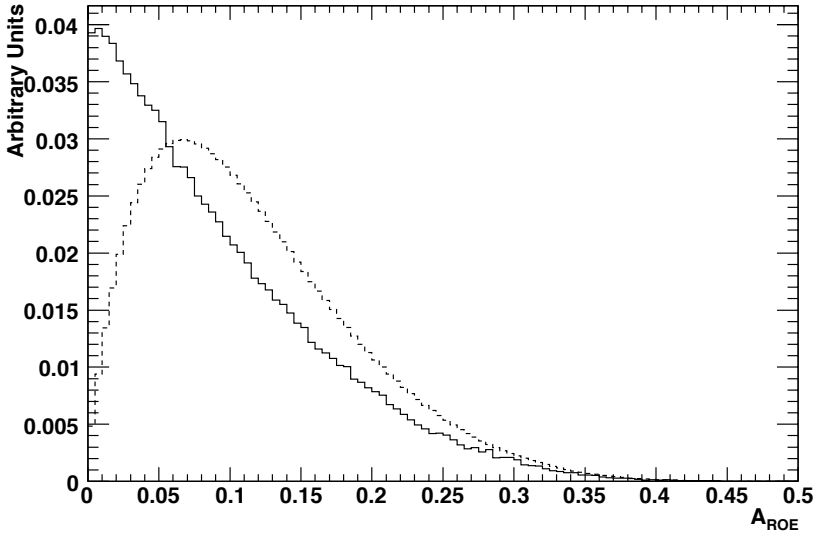


FIGURE A.15: Distributions of A_{ROE}

Appendix B

Signal Monte Carlo Data Event Classification

This appendix contains tables summarising the distributions of hadronic final states in the generated signal MC data. Where a particular final state corresponds to one of the reconstructed modes listed in table 3.1 this is indicated numerically in parentheses after that mode.

B.1 Resonant MC data

B.1.1 $B \to X_s\gamma$

Tables B.1 and B.2 list the distribution of final states for $B^+ \to K^{*+}\gamma$ and $B^0 \to K^{*0}\gamma$ respectively. These distributions are determined by the measured decays of the K^* resonance [14].

Mode	Hadronic Mass $<1.0\,\mathrm{GeV}/c^2$
$B^+ \to K^{*+}(K^0\pi^+)\gamma$	3,827,841
$B^+ \to K^{*+}(K^+\pi^0)\gamma$ (2)	1,918,199
Other 2 body K^{*+}	6,131
Total	5,752,171

TABLE B.1: Distribution of final states in $B^+ \to K^{*+}\gamma$ signal MC data

Mode	Hadronic Mass <1.0 GeV/c^2
$B^0 \to K^{*0}(K^+\pi^-)\gamma$ (1)	3,814,990
$B^0 \to K^{*0}(K^0\pi^0)\gamma$	1,900,490
Other 2 body K^{*0}	11,993
Total	5,727,473

TABLE B.2: Distribution of final states in $B^0 \to K^{*0}\gamma$ signal MC data

Mode	Hadronic Mass <1.0 GeV/c^2
$B^0 \to \omega(\pi^+\pi^-)\gamma$ (1)	15,738
$B^0 \to \omega(\pi^+\pi^-\pi^0)\gamma$ (4)	576,785
$B^0 \to \omega(\pi^+\pi^-\pi^+\pi^-)\gamma$ (5)	596
Other 2 body ω	55,967
Other 3 body ω	914
Total	650,000

TABLE B.3: Distribution of final states in $B^0 \to \omega\gamma$ signal MC data

B.1.2 $B \to X_d\gamma$

Table B.3 lists the distribution of final states in $B^0 \to \omega\gamma$ MC data. This distribution is determined by the measured decays of the ω resonance [14]. The distributions of final states for $B \to \rho\gamma$ modes are not shown as the MC data assumes all $\rho^+ \to \pi^+\pi^0$ and all $\rho^0 \to \pi^+\pi^-$.

B.2 Non-Resonant MC data with Generated Hadronic Mass Following the KN Model

B.2.1 $B \to X_s\gamma$

Tables B.4 and B.5 show the distribution of final states for high mass signal MC data for $B^+ \to X_s^+\gamma$ and $B^0 \to X_s^0\gamma$ decays respectively. The hadronic mass is determined from the KN model and the final state distributions are determined by phase space decays from JETSET [44]. The total number of decays are shown as well as the number with generated hadronic mass within the mass range of reconstructed candidates. The application of weights to these data, as described in section 4.2, is shown in table B.6.

Mode	Hadronic Mass $\geq$1.0 GeV/c^2	Hadronic Mass 1.0-2.0 GeV/c^2
$B^+ \to K^0\pi^+\gamma$	970,100	856,693
$B^+ \to K^+\pi^0\gamma$ (2)	595,196	526,122
$B^+ \to K^+\pi^+\pi^-\gamma$ (3)	1,098,686	915,113
$B^+ \to K^0\pi^+\pi^0\gamma$	1,332,026	1,110,800
$B^+ \to K^+\pi^0\pi^0\gamma$ (8)	271,142	226,589
$B^+ \to K^0\pi^+\pi^-\pi^+\gamma$	391,333	255,914
$B^+ \to K^+\pi^+\pi^-\pi^0\gamma$ (6)	1,004,779	701,015
$B^+ \to K^0\pi^+\pi^0\pi^0\gamma$	346,501	227,603
$B^+ \to K^+\pi^+\pi^-\pi^+\pi^-\gamma$	139,007	61,728
$B^+ \to K^0\pi^+\pi^-\pi^+\pi^0\gamma$	613,763	315,611
$B^+ \to K^+\pi^+\pi^-\pi^0\pi^0\gamma$	397,167	196,662
$B^+ \to K^+\eta\gamma$ (7)	85,037	74,348
$B^+ \to K^0\pi^+\eta\gamma$	63,897	46,157
$B^+ \to K^+\pi^0\eta\gamma$	34,520	24,968
$B^+ \to K^+\pi^+\pi^-\eta\gamma$	35,901	19,172
$B^+ \to K^0\pi^+\pi^0\eta\gamma$	37,407	19,274
$B^+ \to K^+K^-K^+\gamma$	60,945	40,808
$B^+ \to K^0K^+K^-\pi^+\gamma$	55,037	18,265
$B^+ \to K^+K^-K^+\pi^0\gamma$	29,800	10,259
Other 2 body	30	0
Other 3 body	93,615	51,192
Other 4 body	233,179	137,447
Other 5+ body	2,536,332	445,731
Total	10,425,400	6,281,471

TABLE B.4: Distribution of final states in high mass $B^+ \to X_s^+\gamma$ signal MC data. The hadronic mass is determined by the KN model.

B.2.2 $B \to X_d\gamma$

Tables B.7 and B.8 show the distribution of final states for high mass signal MC data for $B^+ \to X_d^+\gamma$ and $B^0 \to X_d^0\gamma$ decays respectively. The hadronic mass is determined from the KN model and the final state distributions are determined by phase space decays from JETSET [44]. The total number of decays are shown as well as the number with generated hadronic mass within the mass range of reconstructed candidates.

Mode	Hadronic Mass $\geq$1.0 GeV/c^2	Hadronic Mass 1.0-2.0 GeV/c^2
$B^0 \to K^+\pi^-\gamma$ (1)	971,350	858,131
$B^0 \to K^0\pi^0\gamma$	594,227	524,904
$B^0 \to K^0\pi^+\pi^-\gamma$	1,104,038	919,640
$B^0 \to K^+\pi^-\pi^0\gamma$ (4)	1,334,309	1,112,561
$B^0 \to K^0\pi^0\pi^0\gamma$	268,178	224,300
$B^0 \to K^+\pi^-\pi^+\pi^-\gamma$ (5)	393,695	256,414
$B^0 \to K^0\pi^+\pi^-\pi^0\gamma$	1,036,398	725,643
$B^0 \to K^+\pi^-\pi^0\pi^0\gamma$ (9)	347,271	228,359
$B^0 \to K^0\pi^+\pi^-\pi^+\pi^-\gamma$	141,785	63,083
$B^0 \to K^+\pi^-\pi^+\pi^-\pi^0\gamma$	569,578	289,191
$B^0 \to K^0\pi^+\pi^-\pi^0\pi^0\gamma$	407,614	203,538
$B^0 \to K^0\eta\gamma$	85,063	74,155
$B^0 \to K^+\pi^-\eta\gamma$	63,890	46,109
$B^0 \to K^0\pi^0\eta\gamma$	34,088	24,576
$B^0 \to K^0\pi^+\pi^-\eta\gamma$	39,168	21,004
$B^0 \to K^+\pi^-\pi^0\eta\gamma$	37,538	19,185
$B^0 \to K^0K^+K^-\gamma$	61,209	41,026
$B^0 \to K^+K^-K^+\pi^-\gamma$	34,781	12,085
$B^0 \to K^0K^+K^-\pi^0\gamma$	29,613	10,024
Other 2 body	29	0
Other 3 body	88,673	47,769
Other 4 body	206,255	110,333
Other 5+ body	2,580,313	465,627
Total	10,429,063	6,277,657

TABLE B.5: Distribution of final states in high mass $B^0 \to X_s^0\gamma$ signal MC data. The hadronic mass is determined by the KN model.

Mode	Weight	Hadronic Mass $\geq$1.0 GeV/c^2
$B^0 \to K^+\pi^-\gamma$ (1)	0.65±0.04	631,378
$B^0 \to K^+\pi^0\gamma$ (2)	0.36±0.06	214,271
$B^0 \to K^+\pi^-\pi^+\gamma$ (3)	1.34±0.11	1,472,239
$B^0 \to K^+\pi^-\pi^0\gamma$ (4)	1.35±0.11	1,801,317
$B^0 \to K^+\pi^-\pi^+\pi^-\gamma$ (5)	0.75±0.27	295,271
$B^0 \to K^+\pi^-\pi^+\pi^0\gamma$ (6)	1.00±0.23	1,004,779
$B^0 \to K^+\eta\gamma$ (7)	1.05±0.41	89,289
2 body no π^0	0.50±0.07	485,050
2 body 1 π^0	0.19±0.12	112,903
3 body no π^0	1.02±0.14	1,126,119
3 body 1 π^0	1.34±0.24	1,784,915
4 body no π^0	2.67±0.96	1,044,859
4 body 1 π^0	1.29±0.61	1,336,953
3/4 body 2 π^0	1.89±1.33	2,330,544
5 body	$1.32\pm^{+1.55}_{-1.32}$	2,994,966
η modes	$0.83\pm^{+1.00}_{-0.83}$	358,122
KKK modes	$0.27\pm^{+0.54}_{-0.27}$	73,274
Unclassified	0.64	3,698,624
Total	1.00	20,854,463

TABLE B.6: Application of weights to high mass $B \to X_s\gamma$ signal MC data.

Mode	$m_X^{true} \geq$1.0 GeV/c^2	1.0$\leq m_X^{true} \leq$2.0 GeV/c^2
$B^+ \to \pi^+\pi^0\gamma$ (2)	110,859	95,550
$B^+ \to \pi^+\pi^-\pi^+\gamma$ (3)	95,291	79,024
$B^+ \to \pi^+\pi^0\pi^0\gamma$ (8)	88,869	73,913
$B^+ \to \pi^+\pi^-\pi^+\pi^0\gamma$ (6)	183,197	142,670
$B^+ \to \pi^+\eta\gamma$ (7)	12,790	11,000
Other 2 body	9,907	6,985
Other 3 body	47,284	34,937
Other 4 body	87,150	58,688
5+ body	566,589	220,478
Total	1,201,936	723,245

TABLE B.7: Distribution of final states in high mass $B^+ \to X_d^+\gamma$ signal MC data. The hadronic mass is determined by the KN model.

Mode	$m_X^{true} \geq 1.0\,\mathrm{GeV}/c^2$	$1.0 \leq m_X^{true} \leq 2.0\,\mathrm{GeV}/c^2$
$B^0 \to \pi^+\pi^-\gamma$ (1)	99,331	85,512
$B^0 \to \pi^+\pi^-\pi^0\gamma$ (4)	197,465	165,434
$B^0 \to \pi^+\pi^-\pi^+\pi^-\gamma$ (6)	44,327	32,459
$B^0 \to \pi^+\pi^-\pi^0\pi^0\gamma$ (9)	131,267	101,180
Other 2 body	44,336	36,704
Other 3 body	53,659	39,488
Other 4 body	62,555	39,280
5+ body	568,768	222,774
Total	1,201,708	722,831

TABLE B.8: Distribution of final states in high mass $B^0 \to X_d^0\gamma$ signal MC data. The hadronic mass is determined by the KN model.

References

[1] *BABAR* Collaboration, B. Aubert *et al.*, "Measurement of $B \to X\gamma$ Decays and Determination of $|V_{td}/V_{ts}|$," *Phys. Rev. Lett.* **102** 161803 (2009), *BABAR* Analysis Document #1340.

[2] N. Cabibbo, "Unitary Symmetry and Leptonic Decays," *Phys. Rev. Lett.* **10** 531 (1963).

[3] M. Kobayashi and T. Maskawa, "CP Violation in the Renormalizable Theory of Weak Interactions," *Prog. Theor. Phys.* **49** 652 (1973).

[4] P. F. Harrison and H. R. Quinn (Ed.), "The *BABAR* Physics Book." Technical Report SLAC-R-504 (1998).

[5] A. Seiden, *Particle Physics a Comprehensive Introduction*, Addison Wesley (2005).

[6] D. Griffiths, *Introduction to Elementary Particles*, John Wiley & Sons, Inc. (1987).

[7] F. Halzen and A. D. Martin, *Quarks & Leptons: An Introductory Course in Modern Particle Physics*, John Wiley & Sons, Inc. (1984).

[8] N. Ellis and C. Dib (Ed.), "2007 CERN-CLAF School of High Energy Physics Proceedings," Technical Report CERN-2008-004 (2008).

[9] T. J. Greenshaw (Ed.), "Proceedings of the School for Experimental High Energy Physics Students held 3 to 15 September 2006," Technical Report RAL-TR-2007-004 (2007).

[10] S. L. Glashow, "Partial Symmetries of Weak Interactions," *Nucl. Phys.* **22** 579 (1961).

[11] A. Salam and J. C. Ward, "Gauge Theory of Elementary Interactions," *Phys. Rev.* **B136** 763 (1964).

[12] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett.* **19** 1264 (1967).

[13] P. W. Higgs, "Broken Symmetries, Massless Particles and Gauge Fields," *Phys. Lett.* **12** 132 (1964).

[14] Particle Data Group, C. Amsler *et al.*, "The Review of Particle Physics," *Phys. Lett.* **B667** 1 (2008).

[15] L. Wolfenstein, "Parametrization of the Kobayashi-Maskawa Matrix," *Phys. Rev. Lett.* **55** 1945 (1983).

[16] A. J. Buras, M. E. Lautenbacher and G. Ostermaier, "Waiting for the top quark mass, $K^+ \to \pi^+ \nu\overline{\nu}$, $B_s^0 - \overline{B}_s^0$ mixing and CP asymmetries in B decays," *Phys. Rev.* **D50** 3433 (1994).

[17] S. Weinberg, *The Quantum Theory of Fields*, Cambridge University Press (1995).

[18] J. H. Christenson *et al.*, "Evidence for the 2π Decay of the K_2^0 Meson," *Phys. Rev. Lett.* **13** 138 (1964).

[19] *BABAR* Collaboration, B. Aubert *et al.*, "Measurement of Time-Dependent CP Asymmetry in $B^0 \to c\bar{c}K^{(*)0}$ Decays," *Phys. Rev.* **D79** 072009 (2009), *BABAR* Analysis Document #1996.

[20] *BABAR* Collaboration, B. Aubert *et al.*, "Observation of CP Violation in $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$," *Phys. Rev. Lett.* **99** 021603 (2007), *BABAR* Analysis Document #1551.

[21] *BABAR* Collaboration, B. Aubert *et al.*, "Evidence for Direct CP Violation from Dalitz-plot analysis of $B^\pm \to K^\pm\pi^\mp\pi^\pm$," *Phys. Rev.* **D78** 012004 (2008), *BABAR* Analysis Document #1181.

[22] A. Ali, H. Asatrian and C. Greub, "Inclusive Decay Rate for $B \to X_d + \gamma$ in Next-to-Leading Logarithmic Order and CP Asymmetry in the Standard Model," *Phys. Lett.* **B429** 87 (1998).

[23] G. Buchalla, A. Buras and M. E. Lautenbacher, "Weak Decays Beyond Leading Logarithms," *Rev. Mod. Phys.* **68** 1125 (1996).

[24] Tobias Hurth, "Present Status of Inclusive Rare B Decays," *Rev. Mod. Phys.* **75** 1159 (2003).

[25] A. J. Buras, "Operator Product Expansion, Renormalisation Group and Weak Decays," *Lect. Notes Phys.* **558** 65 (2000).

[26] A. Ali, E. Lunghi, A. Ya. Parkhomenko, "Implication of the $B \to (\rho, \omega)\gamma$ Branching Ratios for the CKM Phenomenology," *Phys. Lett.* **B595** 323 (2004).

[27] A. L. Kagan and M. Neubert, "QCD Anatomy of $B \to X_s\gamma$ Decays," *Eur. Phys. J.* **C7** 5 (1999).

[28] D. Benson, I.I. Bigi and N. Uraltsev, "On the Photon Energy Moments and their 'Bias' Corrections in $B \to X_s\gamma$," *Nucl. Phys.* **B710** 371 (2005).

[29] B.O. Lange, M. Neubert and G. Paz, "Theory of Inclusive Charmless B Decays and the Extraction of V_{ub}," *Phys. Rev.* **D72** 073006 (2005).

[30] *BABAR* Collaboration, B. Aubert *et al.*, "Measurements of the $B \to X_s\gamma$ Branching Fraction and Photon Spectrum from a Sum of Exclusive Final States," *Phys. Rev.* **D72** 052004 (2005), *BABAR* Analysis Document #768.

[31] O. L. Buchmüller and H. U. Flächer, "Fit to Moments of Inclusive $B \to X_c l\overline{\nu}$ and $B \to X_s\gamma$ Decay Distributions using Heavy Quark Expansions in the Kinetic Scheme," *Phys. Rev.* **D73** 073008 (2006).

[32] CDF Collaboration, A. Abulencia *et al.*, "Observation of B_s^0-$\overline{B}_s^0$ Oscillations." *Phys. Rev. Lett.* **97** 242003 (2006).

[33] *BABAR* Collaboration, B. Aubert *et al.*, "Measurements of Branching Fractions for $B^+ \to \rho^+\gamma$, $B^0 \to \rho^0\gamma$, and $B^0 \to \omega\gamma$." *Phys. Rev.* **D78** 112001 (2008), *BABAR* Analysis Document #1693.

[34] BELLE Collaboration, N. Taniguchi *et al.*, "Measurement of branching fractions, isospin and CP-violating asymmetries for exclusive $b \to d\gamma$ modes." *Phys. Rev. Lett.* **101** 111801 (2008).

[35] *BABAR* Collaboration, D. Boutigny *et al.*, "*BABAR* Technical Design Report." Technical Report SLAC-R-475 (1995).

[36] M. S. Zisman (Ed.), "PEP-II, An Asymmetric B Factory: Conceptual Design Report." Technical Report SLAC-R-418 (1993).

[37] S. Ecklund, C. Field and G. Mazaheri, "A Fast Luminosity Monitor System for PEP II," *Nucl. Instrum. Meth.* **A463** 68 (2001).

[38] *BABAR* Collaboration, B. Aubert *et al.*, "The *BABAR* Detector." *Nucl. Instrum. Meth.* **A479** 1 (2002).

[39] *BABAR* Collaboration, Electron Identification Analysis Working Group, "Cut-Based Electron Identification," *BABAR* Analysis Document #90 (2001).

[40] *BABAR* Collaboration, W. Menges, "The *BABAR* Muon System Upgrade." in "Proceedings of 2005 IEEE Nuclear Science Symposium and Medical Imaging Conference," (2005).

[41] Charge conjugation is implied throughout the subsequent text unless otherwise stated.

[42] *BABAR* Collaboration, B. Aubert *et al.*, "Measurement of the Branching Fraction and Photon Energy Moments of $B \to X_s\gamma$ and $A_{CP}(B \to X_{s+d}\gamma)$," *Phys. Rev. Lett.* **97** 171803 (2006), *BABAR* Analysis Document #323.

[43] *BABAR* Collaboration, B. Aubert *et al.*, "Measurement of the $B \to X_s\gamma$ Branching Fraction and Photon Energy Spectrum using the Recoil Method," *Phys. Rev.* **D77** 051103 (2008), *BABAR* Analysis Document #1609.

[44] T. Sjöstrand, "High Energy Physics Event Genereation with PYTHIA 5.7 and JETSET 7.4," *Comput. Phys. Commun.* **82** 74 (1994).

[45] *BABAR* Collaboration, G. D. McGregor, "B Counting at *BABAR*," Technical Report SLAC-R-912 (2008).

[46] GEANT4 Collaboration, S. Agostinelli *et al.*, "GEANT4: A Simulation Toolkit," *Nucl. Instrum. Meth.* **A506** 250 (2003).

[47] E. Barberio and Z. Was, "PHOTOS: A Univeral Monte Carlo for QED Radiative Corrections." *Comput. Phys. Commun.* **79** 291 (1994).

[48] *BABAR* Collaboration, G. Cowan *et al.*, "Smearing of Monte Carlo Calorimeter Simulation," *BABAR* Analysis Document #1639 (2007).

[49] *BABAR* Collaboration, M. Carpinelli *et al.*, "The *BABAR* Vertexing," *BABAR* Analysis Document #102 (2001).

[50] *BABAR* Collaboration, P. Ongmongkolkul *et al.*, "Particle Identification Using Error Correction Output Code Multiclass Classifier," *BABAR* Analysis Document #2199 (2009).

[51] *BABAR* Collaboration, W. T. Ford, "Choice of Kinematic Variables in B Meson Reconstruction," *BABAR* Analysis Document #53 (2000).

[52] E. Fahri, "Quantum Chromodynamics Test for Jets," *Phys. Rev. Lett.* **39** 1587 (1977).

[53] G.C. Fox and S. Wolfram, "Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes," *Phys. Rev. Lett.* **41** 1581 (1978).

[54] *BABAR* Collaboration, J. Ocariz *et al.*, "Background Fighting in Charmless Two-body Analyses," *BABAR* Analysis Document #346 (2002).

[55] J.D. Bjorken and S.J. Brodsky, "Statistical Model for Electron-Positron Annihilation into Hadrons," *Phys. Rev.* **D1** 1416 (1970).

[56] *BABAR* Collaboration, J. Beringer, "BTagger - A Multivariate Tagging Algorithm with Categories Based on the Physics of the B_{tag} Decay," *BABAR* Analysis Document #317 (2002).

[57] *BABAR* Collaboration, D. Lange *et al.*, "BtgTest and Tag04: Studies Towards an Improved Tagging Algorithm," *BABAR* Analysis Document #730 (2004).

[58] *BABAR* Collaboration, R Bartoldus *et al.*, "Trigger and Filter Documentation for the Run 1 Data," *BABAR* Analysis Document #194 (2002).

[59] *BABAR* Collaboration, B. Aubert *et al.*, "Branching Fractions of $B^+ \to \rho^+\gamma$, $B^0 \to \rho^0\gamma$, and $B^0 \to \omega\gamma$." *Phys. Rev. Lett.* **98** 151802 (2007), *BABAR* Analysis Document #1301.

[60] I. Narsky, "StatPatternRecognition: A C++ Package for Statistical Analysis of High Energy Physics Data," arXiv:physics/0507143 (2005).

[61] W. Verkerke and D. Kirkby, "The RooFit Toolkit for Data Modeling," arXiv:physics/0306116 (2003).

[62] F. James and M. Roos, "'Minuit' A System for Function Minimization and Analysis of the Parameter Errors and Correlations," *Comput. Phys. Commun.* **10** 343 (1975).

[63] L. Lyons, *Statistics for Nuclear and Particle Physicists*, Cambridge University Press (1992).

[64] R. Barlow, *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*, John Wiley & Sons, Ltd. (1989).

[65] ARGUS Collaboration, H. Albrecht *et al.*, "Reconstruction of B Mesons," *Phys. Lett.* **B185** 218 (1987).

[66] J.E. Gaiser, "Hadron Production by e^+e^- Annihilation at Center-of-Mass Energies Between 2.6 and 7.8 GeV," Ph.D Thesis, Technical Report SLAC-255 (1982).

[67] *BABAR* Collaboration, B. Aubert *et al.*, "Measurement of the Branching Fraction and the Time-Dependent CP-Violating Asymmetry for the Decay $B^0 \to K^0_S\pi^0$," *Phys. Rev.* **D71** 111102 (2005), *BABAR* Analysis Document #904.

[68] *BABAR* Collaboration, I.M. Nugent, "Tau31 Tracking Efficiency Study for 2004," *BABAR* Analysis Document #931 (2004).

[69] *BABAR* Collaboration, M. Allen *et al.*, "A Measurement of π^0 Efficiency Using $\tau \to \rho\nu$ and $\tau \to \pi\nu$ Decays," *BABAR* Analysis Document #870 (2004).

[70] *BABAR* Collaboration, D. Payne *et al.*, "Single Photon Efficiency Using $\mu^+\mu^-\gamma$," *BABAR* Analysis Document #1110 (2005).

[71] *BABAR* Collaboration, B. Aubert *et al.*, "Measurement of $B \to K^*\gamma$ Branching Fractions and Charge Asymmetries," *Phys. Rev. Lett.* **88** 101805 (2002), *BABAR* Analysis Document #33.

[72] BELLE Collaboration, A. Abashian *et al.*, "The BELLE Detector," *Nucl. Instrum. Meth.* **A479** 117 (2002).

[73] LHCb Collaboration, A. Augusto Alves *et al.*, "The LHCb Detector at the LHC," *JINST* **3** S08005 (2008).

[74] LHCb Collaboration, G. Pakhlova and I. Belyaev, "Radiative B Decays with LHCb," Technical Report LHCb 2003-090 (2003).

[75] SuperB Collaboration, M. Bona *et al.*, "SuperB: A High-Luminosity Asymmetric e+ e- Super Flavor Factory. Conceptual Design Report," arXiv:0709.0451 [hep-ex] (2007).

[76] BELLE II Collaboration, I. Adachi *et al.*, "sBelle Design Study Report," arXiv:0810.4084 [hep-ex] (2008).

[77] *BABAR* Collaboration, P. del Amo Sanchez *et al.*, "Study of $B \to X\gamma$ Decays and Determination of $|V_{td}/V_{ts}|$." arXiv:1005.4087 [hep-ex] (2010).

"We do not yet know all the basic laws: there is an expanding frontier of ignorance."

Richard Feynman